

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:13:34 ; Search time 163.5 Seconds
(without alignments)
4056.844 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853

Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSESTGFGBERESIL 1715

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8853	100.0	1715	5	Aae25143 Rat ARMS
2	8853	100.0	1715	7	Adc06849 Rat Kidin
3	8639.5	97.6	1762	5	Aau96841 Rat Kidin
4	8639.5	97.6	1762	7	Adc06848 Rat Kidin
5	8526	96.3	1763	5	Aau80244 Rat Kidin
6	8266	93.4	1715	4	Aam39025 Human pol
7	8266	93.4	1715	5	Aam38993 Human pol
8	8266	93.4	1715	5	Aae25144 Human ARM
9	8266	93.4	1715	5	Aau96840 Human kid
10	8266	93.4	1715	6	Aae32128 Human cyt
11	8237	93.0	1753	6	Ada09888 Human rec
12	8212	92.8	1771	7	Adc06847 Human Kid
13	5747	64.9	1142	7	Adc06850 Kidins 22
14	5730.5	64.7	1184	7	Adc06846 Human Kid
15	2961.5	33.5	705	4	Aag75604 Human col
16	2782.5	31.4	551	4	Aae01035 Human dea
17	2408	27.2	543	4	Aab95191 Human pro
18	2376.5	26.8	1498	4	Abb64857 Drosophil
19	2376.5	26.8	1498	5	Aae25146 Fruit fly
20	2260	25.5	513	4	Aae03645 Human ext
21	2150	24.3	624	4	Abg08697 Novel hum
22	1848.5	20.9	1398	5	Aae25145 Worm ARMS
23	662	7.5	131	5	Adk36597 Novel hum
24	614	6.9	120	4	Aam40811 Human pol
25	599.5	6.8	129	4	Aab94785 Human pro

26	558	6.3	1881	7	ADD47763	Add47763 Human pro
27	553	6.2	1880	8	ADR90358	Adr90358 Full leng
28	553	6.2	3897	8	Abm83919	Abm83919 Human dia
29	553	6.2	3924	8	Abm83916	Abm83916 Human dia
30	553	6.2	3955	8	Abm83915	Abm83915 Human dia
31	553	6.2	3985	8	Abm83914	Abm83914 Human dia
32	553	6.2	4016	8	Abm83912	Abm83912 Human dia
33	553	6.2	4072	8	Abm83913	Abm83913 Human dia
34	549.5	6.2	4082	8	Abm83911	Abm83911 Human dia
35	549	6.2	109	3	Aab56277	Aab56277 Human sec
36	542.5	6.1	1955	8	Abm83917	Abm83917 Human dia
37	542	6.1	1954	8	Abm83918	Abm83918 Human dia
38	538.5	6.1	1724	5	AAO20513	AAO20513 Protein o
39	536	6.1	1839	7	ADD27862	Add27862 Human bra
40	534.5	6.0	1872	4	AAM79160	Aam79160 Human pro
41	531.5	6.0	986	4	ABG31785	Abg31785 Human ank
42	531.5	6.0	1139	7	ADC39122	Adc39122 Novel hum
43	531.5	6.0	1250	6	AAE33667	Aae33667 Human str
44	529	6.0	2622	7	ADe55508	Ad555508 Rat Prote
45	527.5	6.0	657	7	ADB64375	Adb64375 Human pro

ALIGNMENTS

RESULT 1
AAE25143
ID AAE25143 standard; protein; 1715 AA.
XX AC AAE25143;
XX AC
DT 30-OCT-2002 (first entry)
XX AC
DE Rat ARMS protein.
XX AC
KW Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW diagnostic; imaging; rat.
XX AC
OS Rattus norvegicus.
XX AC
FH Key Location/Qualifiers
FT Region 17..390
FT FT /note= "Ankyrin repeat"
FT Region 354..493
FT FT /note= "N-terminal region"
FT Domain 496..518
FT FT /note= "Transmembrane domain"
FT Domain 525..546
FT FT /note= "Transmembrane domain"
FT Region 573..638
FT FT /note= "Loop between TM2 and TM3"
FT Domain 661..680
FT FT /note= "Transmembrane domain"
FT Domain 688..710
FT FT /note= "Transmembrane domain"
FT Region 749..854
FT FT /note= "Carboxy tail"
FT Region 940..1060
FT FT /note= "Carboxy tail"
FT Region 1081..1093
FT FT /note= "Polyproline stretch"
FT Domain 1151..1221
FT FT /note= "SAM domain"
FT Binding-site 1713..1715
FT FT /note= "PDZ binding motif"
XX WO200250273-A2.
XX 27-JUN-2002.
XX 19-DEC-2001; 2001WO-US048603.

Db 1261 ELSSQTPVTLNFSFEELNTGLDEGAPRHNSLWSQSTRTPSLSLNSQSSIEISKLT 1320
Qy 1321 DKVQAEYRDAYREYIAQMSQLEGTSSTISGRSSPHSTYYIGQSSGGSIHSTLQERG 1380
Db 1321 DKVQAEYRDAYREYIAQMSQLEGTSSTISGRSSPHSTYYIGQSSGGSIHSTLQERG 1380
Qy 1381 KEGELKQEDGKRSFLMKRGDIVDYSSSGVSTNEASPLDPTITERDEKSDQSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGKRSFLMKRGDIVDYSSSGVSTNEASPLDPTITERDEKSDQSGSKLLPGKKS 1440
Qy 1441 SERPSLFTDLKLGKGLRYQKLPSEDESGTRGVQITPHCSKMWIRTKRLKAKQRECAPS 1500
Db 1441 SERPSLFTDLKLGKGLRYQKLPSEDESGTRGVQITPHCSKMWIRTKRLKAKQRECAPS 1500
Qy 1501 QEHSAPRIFIRIKAYEYLSALLDKKSDSGVRSNESHSHLHNEAADSQLEKANLI 1560
Db 1501 QEHSAPRIFIRIKAYEYLSALLDKKSDSGVRSNESHSHLHNEAADSQLEKANLI 1560
Qy 1561 ELEDGHSGRGMPHSLSGLODPTIARMSICSEDKSPSCSLIASSPESWPACOKAYN 1620
Db 1561 ELEDGHSGRGMPHSLSGLODPTIARMSICSEDKSPSCSLIASSPESWPACOKAYN 1620
Qy 1621 LNRTPTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPGSPNPTAVQENLKSMAHK 1680
Db 1621 LNRTPTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPGSPNPTAVQENLKSMAHK 1680
Qy 1681 RSQRSSYTRLKDASELHAASSESTGFERESIL 1715
Db 1681 RSQRSSYTRLKDASELHAASSESTGFERESIL 1715

RESULT 3

ID AAU96841 standard; protein; 1762 AA.
AC AAU96841;

XX
DT 09-SEP-2004 (revised)
DT 30-JUL-2002 (first entry)
XX
DE Rat kidins220 protein from Genbank AAG35185.
XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX

OS Unidentified.

XX WO200220786-A2.
XX

PN 14-MAR-2002.

XX 06-SEP-2001; 2001WO-GB003977.

XX 06-SEP-2000; 2000US-0230449P.

XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

XX Schiavo G, Iglesias T;

XX WPI; 2002-371879/40.

XX Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,
PT useful for identifying modulators useful in treating cancer and
PT neurodegenerative diseases.
XX

PS Disclosure; Fig 18; 228pp; English.

XX The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their fragment,
CC variant or fusion provided that the protein is not the polypeptide

CC encoded by GenBank Accession No. AB033076. Also included are a
CC polynucleotide encoding Kidins220 provided that polynucleotide is not any
CC one of the clones corresponding to the 61 GenBank Accession Nos. given in
CC the specification; an expression/replicable vector comprising the
CC polynucleotide; a recombinant host cell containing the polynucleotide or
CC vector; an anti-Kidins220 antibody (used in the preparation of Kidins220)
CC ; an agent (A1) which modulates activity of Kidins220, protein kinase D
CC (PKD) or interaction between PKD and Kidins220; a mutant animal
CC transgenic for kidins220; the use of an agent capable of detecting the
CC expression of Kidins220 gene in manufacture of a diagnostic reagent for
CC diagnosing or prognosing cancer or for monitoring the progression of
CC cancer in a patient; a compound comprising a moiety that selectively
CC binds to Kidins220 or its variant and another moiety; and a nucleic acid
CC molecule encoding the compound. The antibody is useful for modulating the
CC activity of Kidins220, where the antibody prevents a region of Kidins220
CC interacting from another protein, or reduces the ability of Kidins220 to
CC bind to adenosine triphosphate (ATP). The kidins220 polynucleotide is
CC useful for identifying an agent which modulates the promoter activity of
CC the polynucleotide, and Kidins220 is useful for identifying an agent
CC which modulates activity of Kidins220, protein kinase D (PKD) and the
CC interaction between PKD with Kidins220. Kidins220 or its polynucleotide
CC is useful in medicine, e.g. using gene therapy or for diagnosing or
CC monitoring progression of cancer in a patient. The agent is useful in
CC manufacture of medicament, for use in treatment of neurodegenerative
CC disease, and the agent which inhibits function of Kidins220 gene or its
CC product is useful for treating cancer which is glioblastoma multiforme or
CC prostate cancer. The compound is useful for imaging cancer in an
CC individual, and for diagnosing or prognosing, and also for treating an
CC individual. The present sequence represents Rat kidins220 protein from
CC Genbank accession number AAG35185

CC Revised record issued on 09-SEP-2004 : Correction to Organism field

XX Sequence 1762 AA;
SQ

Query Match 97.6%; Score 8639.5; DB 5; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

Qy 1 MSVLISQSVINVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGVVEIVKELLKN 60
Db 1 MSVLISQSVINVEEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGVVEIVKELLKN 60

Qy 61 GANCNLEDJNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACVKGRTDVV 120
Db 61 GANCNLEDJNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACVKGRTDVV 120

Qy 121 ELLLSHGANSVTLGQYSVYPIIWAAGRGHADIHVHLLQNGAKVNCSDKYGTTPPLVWAAR 180
Db 121 ELLLSHGANSVTLGQYSVYPIIWAAGRGHADIHVHLLQNGAKVNCSDKYGTTPPLVWAAR 179

Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILKNPNVNLTKDQGNITA 240
Db 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILKNPNVNLTKDQGNITA 239

Qy 241 LMTASKEGHEIIVQDILLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKYADIDIRG 300
Db 240 LMTASKEGHEIIVQDILLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKYADIDIRG 299

Qy 301 QDNKTALYWAWEKGNATMVDRDILQCNPDTEICTKOGTEPLIKATMRNIEVVELLDKGA 360
Db 300 QDNKTALYWAWEKGNATMVDRDILQCNPDTEICTKOGTEPLIKATMRNIEVVELLDKGA 359

Qy 361 KVSADVKKGDTPLHVAIRGSRRLAELLRLNPKDGLLRPNKAGETPNIDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHVAIRGSRRLAELLRLNPKDGLLRPNKAGETPNIDCSHOKSIL 419

Qy 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKGSFLLKKL 480
Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKGSFLLKKL 479

Qy 481 EDEMKTFAQOQTEPLFQFSLWLVFLTLCCGGLGVFAFPDVTNLAIALISLSFLAIYIF 540
Db 481 EDEMKTFAQOQTEPLFQFSLWLVFLTLCCGGLGVFAFPDVTNLAIALISLSFLAIYIF 540

Db 480 EDEMTFAGQQTPELPQFSWLIIVFTLLCGGLGVLPAPTVDNTNLAIASISFLALIYIF 539
Qy 541 FTVIYFGRRGESNMAWALSTRLARHIGYLELLFKLMFVNPPPELQPTTKALPVRELF 600
Db 540 FTVIYFGRRGESNMAWALSTRLARHIGYLELLFKLMFVNPPPELQPTTKALPVRELF 599
Qy 601 TDYRLSSVGGETSIAEMIATLSDACEREPGLATRLFRVFRTRTESQKKKKWKTCCCLPS 660
Db 600 TDYRLSSVGGETSIAEMIATLSDACEREPGLATRLFRVFRTRTESQKKKKWKTCCCLPS 659
Qy 661 FVIFLIVGCIAGITLAIARVDPKHLTVNAILISIASVVGCLAPVLCNRTWQVLDL 720
Db 660 FVIFLIVGCIAGITLAIARVDPKHLTVNAILISIASVVGCLAPVLCNRTWQVLDL 719
Qy 721 NSQRRLHSAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQRLVVIIDGLDAC 780
Db 720 NSQRRLHSAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQRLVVIIDGLDAC 779
Qy 781 EODKVLQMLDTRVRLFSKGPPIAIPASDPHIIKAINONLNSVLKSDNSINGHDYMRNIVH 840
Db 780 EODKVLQMLDTRVRLFSKGPPIAIPASDPHIIKAINONLNSVLKSDNSINGHDYMRNIVH 839
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTGTGTOEDTDRVSONSLGEMTKLSKTA 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTGTGTOEDTDRVSONSLGEMTKLSKTA 899
Qy 901 RRDYRRQMQRTITRQMSFDLTKLLVTEDEWFSDISPOTMRLLNIVSVTGRLLRANQIT 960
Db 900 RRDYRRQMQRTITRQMSFDLTKLLVTEDEWFSDISPOTMRLLNIVSVTGRLLRANQIT 959
Qy 961 FNVRLASWINLTQMPVYRTSWLILYLBETEGLPDQMTLTKMYERISKNIPPTTKDVRPL 1020
Db 960 FNVRLASWINLTQMPVYRTSWLILYLBETEGLPDQMTLTKMYERISKNIPPTTKDVRPL 1019
Qy 1021 EIDGDIRNFVFLSRPVLVARVKTFLPCTVNLDPKREIIADVRAAREQINIGGLAY 1080
Db 1020 EIDGDIRNFVFLSRPVLVARVKTFLPCTVNLDPKREIIADVRAAREQINIGGLAY 1079
Qy 1081 PPLPLHEGPPPPGSGYSPASVCSASFPNGVSPQPHSSVYSGLSGSPQHPFYNR - 1139
Db 1080 PPLPLHEGPPPPGSGYSPASVCSASFPNGVSPQPHSSVYSGLSGSPQHPFYNR 1139
Qy 1140 -----AAVPATGSSLL 1151
Db 1140 PFAPLYTPRYPGSQHLSRSVKTSLPRDQNGPLPCDSGFNKQRAAPVATGSSLL 1199
Qy 1152 SSMVTVDVCEKRLQIEGLDQNMPOYCTTIKANINGRVLSCQNDIKKEMANFGDWH 1211
Db 1200 SSMVTVDVCEKRLQIEGLDQNMPOYCTTIKANINGRVLSCQNDIKKEMANFGDWH 1259
Qy 1212 LFRSMVLEMRVSVQVVPEDPRFLNENSSAPVPHGESARRSHTELPTELSSQTPYTLN 1271
Db 1260 LFRSMVLEMRVSVQVVPEDPRFLNENSSAPVPHGESARRSHTELPTELSSQTPYTLN 1319
Qy 1272 PSPELNLGLDEGAPRHSNLWSQSTRTPSLSSLNSQSSIBISKLTDKVQAEYRDAY 1331
Db 1320 PSPELNLGLDEGAPRHSNLWSQSTRTPSLSSLNSQSSIBISKLTDKVQAEYRDAY 1379
Qy 1332 REYIAQMSOLEGCGTSISGRSSPHSTYIYGQSSSGSHSTLEQERKEGELKQEDGR 1391
Db 1380 REYIAQMSOLEGCGTSISGRSSPHSTYIYGQSSSGSHSTLEQERKEGELKQEDGR 1439
Qy 1392 KSFLMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSLPGKKSERPSLFTDL 1451
Db 1440 KSFLMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSLPGKKSERPSLFTDL 1499
Qy 1452 KLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIIRTKLAK-QRECAAPQESAEPIRT 1510
Db 1500 KLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIIRTKLAK-QRECAAPQESAEPIRT 1557
Qy 1511 FIKAKYLSALLDKQSSDSGVRSNESPNSHLSHNEAADSQLEKANLIELEDEGHSGK 1570
Db 1558 FIKAKYLSALLDKQSSDSGVRSNESPNSHLSHNEAADSQLEKANLIELEDEGHSGK 1617

Qy 1571 RGMPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESWPACQKAYNLNRTSTVTL 1630
Db 1618 RGMPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESWPACQKAYNLNRTSTVTL 1677
Qy 1631 NNNTAPTNRANQNDFEIEGRETQSQVILRPGSPNPTAVQENLKSMAHKSQRSSYTRL 1690
Db 1678 NNNTAPTNRANQNDFEIEGRETQSQVILRPGSPNPTAVQENLKSMAHKSQRSSYTRL 1737
Qy 1691 SKDASELHAASSESTGFGEERESIL 1715
Db 1738 SKDASELHAASSESTGFGEERESIL 1762

RESULT 4

ADC06848
ID ADC06848 standard; protein; 1762 AA.

XX
AC ADC06848;

XX
DT 18-DEC-2003 (first entry)

XX
DE Rat Kidins220Pc protein AF239045.

XX
KW cytotstatic; prostate cancer; breast; gene therapy; transgenic; rat;
Kidins220Pc; kinase D-interacting substrate of 220KDa.

XX
OS Rattus sp.

XX
FN WO2003064599-A2.

XX
PD 07-AUG-2003.

XX
PF 24-JAN-2003; 2003WO-US001943.

XX
PR 25-JAN-2002; 2002US-00054935.

XX
PR 14-FEB-2002; 2002US-0356130P.

XX
PR 22-MAR-2002; 2002US-00102946.

XX
PR 08-APR-2002; 2002US-00117229.

XX
PR 14-MAY-2002; 2002US-00144198.

XX
PR 19-JUL-2002; 2002US-00197824.

XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX
PI Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;

XX
DR WPI; 2003-679495/64.

XX
PT New isolated polynucleotide related to cancer genes, useful for
detecting, diagnosing, staging, monitoring, prognosticating, preventing
or treating cancers, e.g. breast and prostate cancers.

XX
PS Disclosure; Fig 20; 128pp; English.

XX
CC The invention relates to a novel isolated polynucleotide comprising a
differentially-regulated mammalian cancer gene. The polynucleotides of
the invention demonstrate cytostatic activity and are differentially
expressed in prostate cancer. The polynucleotide, polypeptides and
methods of the invention may be useful for detecting, diagnosing,
staging, monitoring, prognosticating, preventing or treating cancers,
particularly breast and prostate cancers. Furthermore, the invention may
be utilized during gene therapy procedures or in the production of
transgenic animals. The current sequence is that of the prostate cancer-
related protein of the invention. The current sequence is not fully
legible within the specification; the complete sequence was obtained from
GenBank.

XX
SQ Sequence 1762 AA;

Query Match 97.6%; Score 8639.5; DB 7; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

QY 1 MSLVLSQSVINYVEENIPALKALLSKCKDVDERNECGOTPLMLAAEQGNVEIVKELLKN 60
DB 1 MSLVLSQSVINYVEENIPALKALLSKCKDVDERNECGOTPLMLAAEQGNVEIVKELLKN 60
QY 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDGGGTALMMWACYKGRTDVV 120
DB 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDGGGTALMMWACYKGRTDVV 120
QY 121 ELLSHGANPSVTGLOYSVPIIWAAGRHADIIVHLLQNGAKVNSDKYGTTPVWAAR 180
DB 121 ELLSHGANPSVTGL-YSVYPIIWAAGRHADIIVHLLQNGAKVNSDKYGTTPVWAAR 179
QY 181 KGHLECVKHLANGADVDOGANSMTALIVAVKGGYTSQVKEILKRNPNVNLTDKDGNTA 240
DB 180 KGHLECVKHLANGADVDOGANSMTALIVAVKGGYTSQVKEILKRNPNVNLTDKDGNTA 239
QY 241 LMTASKEGHIIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
DB 240 LMTASKEGHIIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
QY 301 QDNKTALYWAVERGNATWDRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 360
DB 300 QDNKTALYWAVERGNATWDRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 359
QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLNPNKDGRLLYRPNKAGETPPYINIDCSHOKSIL 420
DB 360 KVSADVKKGDTPLHVAIRGSRRLAELLNPNKDGRLLYRPNKAGETPPYINIDCSHOKSIL 419
QY 421 TQIFGARHLSPTTGDGMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSKFLKKL 480
DB 420 TQIFGARHLSPTTGDGMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSKFLKKL 479
QY 481 EDEWKTFAQOOTEPLFOFSLIIVFLTLILCGGLGFVAFVDTNLAIAISLSFLAIYIF 540
DB 480 EDEWKTFAQOOTEPLFOFSLIIVFLTLILCGGLGFVAFVDTNLAIAISLSFLAIYIF 539
QY 541 FIVYFGREGESNNWAWLSRFLARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLL 600
DB 540 FIVYFGREGESNNWAWLSRFLARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLL 599
QY 601 TDYNNLSSVGGETSLEAMIALSDACEREFGLATRLFRVFRTEESQKKKKWKTCCPLS 660
DB 600 TDYNNLSSVGGETSLEAMIALSDACEREFGLATRLFRVFRTEESQKKKKWKTCCPLS 659
QY 661 FVIFLFTVGCIIAGITTLAIIFRVDPKHLTVNAILISIASVVGGLAFVLCNCTWQVLDLSL 720
DB 660 FVIFLFTVGCIIAGITTLAIIFRVDPKHLTVNAILISIASVVGGLAFVLCNCTWQVLDLSL 719
QY 721 NSQRKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTIDSFQNTQRLVWIIDGLDAC 780
DB 720 NSQRKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTIDSFQNTQRLVWIIDGLDAC 779
QY 781 EQDKVQLMDTVRVLFSKGPPIAIFASDPHIIIIKAINONLSNVLSDNSNINGHDMRNIVH 840
DB 780 EQDKVQLMDTVRVLFSKGPPIAIFASDPHIIIIKAINONLSNVLSDNSNINGHDMRNIVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSIGEMTKLGSKTALN 900
DB 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSIGEMTKLGSKTALN 899
QY 901 RRDYTRRQWORTITRQMSFDLTKLLVTEWDFSDISPTQWRRLLNIVSVTGRLLRANQIT 960
DB 900 RRDYTRRQWORTITRQMSFDLTKLLVTEWDFSDISPTQWRRLLNIVSVTGRLLRANQIT 959
QY 961 FNWDRLASWINLTQWVPYRTSWLILYEETEGLPDQMTLKTMYERISKNIPPTTKDVEPFL 1020
DB 960 FNWDRLASWINLTQWVPYRTSWLILYEETEGLPDQMTLKTMYERISKNIPPTTKDVEPFL 1019
QY 1021 EIDGDIRNFVFLSSRTPVLVARVKTFLCTVNLDPKLREIADVRAAREQINIGGLAY 1080
DB 1020 EIDGDIRNFVFLSSRTPVLVARVKTFLCTVNLDPKLREIADVRAAREQINIGGLAY 1079
QY 1081 PPLPLHEGPPRPPSGYSPASVCSASFNGPFGGVSPQPHSSYYSGLSQPHFFYNNR- 1139

DB 1080 PPLPLHEGPPRPPSGYSPASVCSASFNGPFGGVSPQPHSSYYSGLSQPHFFYNNR- 1139
QY 1140 -----AAVPATGSSLLL 1151
DB 1140 FFAPYLYTPRYYPGSGQHLISRSSVKTSLPRDQNNGLPCDSGFNQRQAAVPATGSSLLL 1199
QY 1152 SSMVTVDVCEKURQIEGLDQNNMPOYCTTIKKANINGRVLSCNIDELKCKEMAMNFGDWH 1211
DB 1200 SSMVTVDVCEKURQIEGLDQNNMPOYCTTIKKANINGRVLSCNIDELKCKEMAMNFGDWH 1259
QY 1212 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTILN 1271
DB 1260 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTILN 1319
QY 1272 FSFEELNTLGLDEGAPRHNSLWSQOTRTPSLNSODSSIEISKLTKVQAEVRDAY 1331
DB 1320 FSFEELNTLGLDEGAPRHNSLWSQOTRTPSLNSODSSIEISKLTKVQAEVRDAY 1379
QY 1332 REVIAQMSOLEGTGSGSTISGRSSPHSTVYIIGQSSGSGSIHSTLEQERKEGELKQEDGR 1391
DB 1380 REVIAQMSOLEGTGSGSTISGRSSPHSTVYIIGQSSGSGSIHSTLEQERKEGELKQEDGR 1439
QY 1392 KSFLMKRGDVIDYSSSGVSTNEASPLDPIETEDEKSDQSGSKLLPKKKSERPSLPQTDL 1451
DB 1440 KSFLMKRGDVIDYSSSGVSTNEASPLDPIETEDEKSDQSGSKLLPKKKSERPSLPQTDL 1499
QY 1452 KLKGGGLRYOKLPDSDESGTGRVQITPHCSKMIRTKLAK- QRECAPSQEHSASPRT 1510
DB 1500 KLKGGGLRYOKLPDSDESGTGRVQITPHCSKMIRTKLAK- QRECAPSQEHSASPRT 1557
QY 1511 FIKAKYVLSDDLDDKDDSDSGVRSNESPNSHLSHNEAADDSDQLEKANLIELEDEGHSGK 1570
DB 1558 FIKAKYVLSDDLDDKDDSDSGVRSNESPNSHLSHNEAADDSDQLEKANLIELEDEGHSGK 1617
QY 1571 RGMPHSLSGLOPPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTPTVTIL 1630
DB 1618 RGMPHSLSGLOPPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTPTVTIL 1677
QY 1631 NNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPTAVQENLKSMAHKRSORSYTRL 1690
DB 1678 NNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPTAVQENLKSMAHKRSORSYTRL 1737
QY 1691 SKDASELHAASSESTGFEERESIL 1715
DB 1738 SKDASELHAASSESTGFEERESIL 1762
RESULT 5
AAU80244
ID AAU80244 standard; protein; 1763 AA.
XX AAU80244;
AC AAU80244;
XX XX
DT 09-SEP-2004 (revised)
DT 30-JUL-2002 (first entry)
XX XX
DE Rat kidins220 protein.
XX XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
OS Rattus sp.
OS Unidentified.
XX XX
PN WO200220786-A2.
XX XX
PD 14-MAR-2002.
XX XX
PF 06-SEP-2001; 2001WO-GB003977.
XX XX
PR 06-SEP-2000; 2000US-0230449P.

XX PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 XX PI Schiavo G, Iglesias T;
 XX DR WPI: 2002-371879/40.
 XX DR N-PSDB; ABK51214.
 XX PT Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,
 XX PT useful for identifying modulators useful in treating cancer and
 XX PT neurodegenerative diseases.
 XX PS Claim 33; Fig 1; 228pp; English.

XX CC The invention relates to polypeptide comprising a kinase D interacting
 CC substrate of 220 kDa (Kidins220) from rat and human or their fragment,
 CC variant or fusion provided that the protein is not the polypeptide
 CC encoded by GenBank Accession No. AB033076. Also included are a
 CC polynucleotide encoding Kidins220 provided that polynucleotide is not any
 CC one of the clones corresponding to the 61 GenBank Accession Nos. given in
 CC the specification; an expression/replicable vector comprising the
 CC polynucleotide; a recombinant host cell containing the polynucleotide or
 CC vector; an anti-Kidins220 antibody (used in the preparation of Kidins220)
 CC ; an agent (AI) which modulates activity of Kidins220, protein kinase D
 CC (PKD) or interaction between PKD and Kidins220; a mutant animal
 CC transgenic for Kidins220; the use of an agent capable of detecting the
 CC expression of Kidins220 gene in manufacture of a diagnostic reagent for
 CC diagnosing or prognosing cancer or for monitoring the progression of
 CC cancer in a patient; a compound comprising a moiety that selectively
 CC binds to Kidins220 or its variant and another moiety; and a nucleic acid
 CC molecule encoding the compound. The antibody is useful for modulating the
 CC activity of Kidins220, where the antibody prevents a region of Kidins220
 CC interacting from another protein, or reduces the ability of Kidins220 to
 CC bind to adenosine triphosphate (ATP). The Kidins220 polynucleotide is
 CC useful for identifying an agent which modulates the promoter activity of
 CC the polynucleotide, and Kidins220 is useful for identifying an agent
 CC which modulates activity of Kidins220, protein kinase D (PKD) and the
 CC interaction between PKD and Kidins220. Kidins220 or its polynucleotide
 CC is useful in medicine, e.g. using gene therapy or for diagnosing or
 CC monitoring progression of cancer in a patient. The agent is useful in
 CC manufacture of medicament, for use in treatment of neurodegenerative
 CC disease, and the agent which inhibits function of Kidins220 gene or its
 CC product is useful for treating cancer which is glioblastoma multiforme or
 CC prostate cancer. The compound is useful for imaging cancer in an
 CC individual, and for diagnosing or prognosing, and also for treating an
 CC individual. The present sequence represents Rat kidins220

Revised record issued on 09-SEP-2004 : Correction to Organism field

Sequence 1763 AA;

Query Match 96.3%; Score 8526; DB 5; Length 1763;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1671; Conservative 10; Mismatches 32; Indels 52; Gaps 6;

Qy 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECCQTPMLAAEQGNVEIVKELKN 60
 Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECCQTPMLAAEQGNVEIVKELKN 60
 Qy 61 GANCNLELDNWTALISASKEGHIHVEELKSGASLEHRDMGGWTALMWACYKGRITDV 120
 Db 61 GANCNLELDNWTALISASKEGHIHVEELKSGASLEHRDMGGWTALMWACYKGRITDV 120
 Qy 121 ELLLSHGANSPTGLQYSVPIIWAAGRGHADI VHLHLLONGAKVNCSDKYGTTPLVWAAR 180
 Db 121 ELLLSHGANSPTGLQYSVPIIWAAGRGHADI VHLHLLONGAKVNCSDKYGTTPLVWAAR 180
 Qy 181 KGHLECVKHLAMGADVDEGANSTALI VAVKGYTQSVKEILKRNPNVNLTKDGNTA 240
 Db 181 KGHLECVKHLAMGADVDEGANSTALI VAVKGYTQSVKEILKRNPNVNLTKDGNTA 240
 Qy 241 LMIASKEGHIHVEQLLDAGTYVNI PDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
 Db 241 LMIASKEGHIHVEQLLDAGTYVNI PDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Db 241 LMIASKEGHIHVEQLLDAGTYVNI PDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
 Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGETPLIKATQWRNIEVVVELLDKGA 360
 Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGETPLIKATQWRNIEVVVELLDKGA 360
 Qy 361 KVSADVKKGDTPLHVAIRGRSRRRLAELLLRNPKOGRLLYRPNKAGETPNIDCSHQKSL 420
 Db 361 KVSADVKKGDTPLHVAIRGRSRRRLAELLLRNPKOGRLLYRPNKAGETPNIDCSHQKSL 420
 Qy 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQSGSKSLKKL 480
 Db 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQSGSKSLKKL 480
 Qy 481 EDEMKTFAQOQTEPLFQFSWLIIVFLTLCCGGLGVPAFPVDTNLAIASISFLALIYIF 540
 Db 481 EDEMKTFAQOQTEPLFQFSWLIIVFLTLCCGGLGVPAFPVDTNLAIASISFLALIYIF 540
 Qy 541 FIVYFGRREGESGNNAWALSTRLARHIGYLELLFKLMFVNPPPELPEQTTKALPVRLF 600
 Db 541 FIVYFGRREGESGNNAWALSTRLARHIGYLELLFKLMFVNPPPELPELAEQTTKALPVRLF 600
 Qy 601 TDYNRLLSSVGGETSLAEWIALTSDACEREFGLATRLFRVPRTEESQKKKWKTCCLPS 660
 Db 601 TDYNRLLSSVGGETSLAEWIALTSDACEREFGLATRLFRVPRTEESQKKKWKTCCLPS 660
 Qy 661 FVIFLFIIVGCIAGITTLAI FRVDPKHLTVNAIISIASVVGGLAFVLCRTWQVDSLL 720
 Db 661 FVIFLFIIVGCIAGITTLAI FRVDPKHLTVNAIISIASVVGGLAFVLCRTWQVDSLL 720
 Qy 721 NSQKRLHSAASKLHKLKSEGFMKVLKCEVELMARMKTI DSFTQNTQRLVVIIDGLDAC 780
 Db 721 NSQKRLHSAASKLHKLKSEGFMKVLKCEVELMARMKTI DSFTQNTQRLVVIIDGLDAC 780
 Qy 781 EODKVLQMLDTRVRLFSKGPPIAIFASDPHII IKAQNQLNSVLKSDNSINGHDYMRNVH 840
 Db 781 EODKVLQMLDTRVRLFSKGPPIAIFASDPHII IKAQNQLNSVLKSDNSINGHDYMRNVH 840
 Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSONSLGEMTKLGSKTALN 900
 Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSONSLGEMTKLGSKTALN 900
 Qy 901 RRDYTRRRQMORTITRQMSFDLTLLVTEDFSDISQPTMERLLNIVSVTGRLLRANQIT 960
 Db 901 RRDYTRRRQMORTITRQMSFDLTLLVTEDFSDISQPTMERLLNIVSVTGRLLRANQIT 960
 Qy 961 FNWDRLASWINLTQOWPYRTSWLIYLEETEGLPDQMTLKIYERISKNIPTTKDVEPLL 1020
 Db 961 FNWDRLASWINLTQOWPYRTSWLIYLEETEGLPDQMTLKIYERISKNIPTTKDVEPLL 1020
 Qy 1021 BIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLR EIIADVRAAREQINIGGLAY 1080
 Db 1021 BIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLR EIIADVRAAREQINIGGLAY 1080
 Qy 1081 PPLPLEHGPPRPSPGYSQASVCSASFGNGFPFGVSPQPHSSYSGSLSGPQHPFYNR- 1139
 Db 1081 PPLPLEHGPPRPSPGYSQASVCSASFGNGFPFGVSPQPHSSYSGSLSGPQHPFYNR 1140
 Qy 1140 -----AAVP-----ATGSLLLSMTVD 1157
 Db 1141 FFAPLYTPRYPGGSQHLISRSVSKASLPDRNNGLP CDSGFNKRQGRSPCHRQLTVT 1200
 Qy 1158 WV-----CEKLRQIEGLDQNMPOYCTTIKKANINGRVLSCQNIDELKEMWAMNFGDWH 1211
 Db 1201 EFNDRGRCCKLRQIEGLDQNMPOYCTTIKKANINGRVLSCQNIDELKEMWAMNFGDWH 1260
 Qy 1212 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTTELPLTELSSQTPYTLN 1271
 Db 1261 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTTELPLTELSSQTPYTLN 1320
 Qy 1272 FSPFELNTGLDEGAPRHSNLSWOSQTRRTPSLSSLSNQSODSSIIISKLTQKVAEYRDAY 1331
 Db 1321 FSPFELNTGLDEGAPRHSNLSWOSQTRRTPSLSSLSNQSODSSIIISKLTQKVAEYRDAY 1380

QY 1332 REYIAQMSQLEGGTSGSTISGRSPHSTYYIGOSSSGSIHSTLEQERKEGELKQEDGR 1391
 Db 1381 REYIAQMSQLEGGTSGSTISGRSPHSTYYIGOSSSGSIHSTLEQERKEGELKQEDGR 1440
 QY 1392 KSFLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSERPFLFOTDL 1451
 Db 1441 KSFLMKRGDVIDYSSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSERPFLFOTDL 1500
 QY 1452 KLKGGGLRYQKLPSDEDESGTGRVQIIPHCCKMIRTKRLKAK-ORECASPOEHSABPIRT 1510
 Db 1501 KLKGGGLRYQKLPSDEDESGTGRVQIIPHCCKMIRTKRLKAK-ORECASPOEHSABPIRT 1558
 QY 1511 FIKAKEVLSALLDKDSSGVSNESSPNHSHLNEAADDSOLEKANLIELEDEGHSGK 1570
 Db 1559 FIKAKEVLSALLDKDSSGVSNESSPNHSHLNEAADDSOLEKANLIELEDEGHSGK 1618
 QY 1571 RGMPHSLSGIQLDPIIARMSICSEDKSPSCSLIASSPEESWPACQKAYNLNRTPTVTIL 1630
 Db 1619 RGMPHSLSGIQLDPIIARMSICSEDKSPSCSLIASSPEESWPACQKAYNLNRTPTVTIL 1678
 QY 1631 NNNTAPTNRANQNPFDETEGIRETSQVILRPGSPNPPTAVONENLKSMAHRSQSSVTRL 1690
 Db 1679 NNNTAPTNRANQNPFDETEGIRETSQVILRPGSPNPPTAVONENLKSMAHRSQSSVTRL 1738
 QY 1691 SKDASELHAASSESTGFGEERESIL 1715
 Db 1739 SKDASELHAASSESTGFGEERESIL 1763

RESULT 6
 AAM39025
 ID AAM39025 standard; protein; 1715 AA.
 AC AAM39025;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2170.
 XX
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 22-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AA158181.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 4; SEQ ID NO 2170; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nontropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 1715 AA;
 Query Match 93.4%; Score 8266; DB 4; Length 1715;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
 QY 1 MSVLISQSVINYVERENIPALKALLEKCKDVERNECGQTPLMLAAEQGNVEIVKELLKN 60
 Db 1 MSVLISQSVINYVERENIPALKALLEKCKDVERNECGQTPLMLAAEQGNVEIVKELLKN 60
 QY 61 GANCNLEDLDNWTALISASKEGHIHVEELLESGASIEHRDMGGWTALMWACYKGRDVV 120
 Db 61 GANCNLEDLDNWTALISASKEGHIHVEELLESGASIEHRDMGGWTALMWACYKGRDVV 120
 QY 121 ELLLSHGANGPSVTGLQYSVYPIIWAAGRGHADIVHLLQNGAKVNGSDKYTTPLVWAAR 180
 Db 121 ELLLSHGANGPSVTGLQYSVYPIIWAAGRGHADIVHLLQNGAKVNGSDKYTTPLVWAAR 180
 QY 181 KGHLECVKHLANGADVDEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240
 Db 181 KGHLECVKHLANGADVDEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240
 QY 241 LMTASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
 Db 241 LMTASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
 QY 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 Db 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPPYINICSHQKSIL 420
 Db 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPPYINICSHQKSIL 420
 QY 421 TQIFGARHLSPTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 480
 Db 421 TQIFGARHLSPTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 480
 QY 481 EDEMTKTFAGQOETPLFQFQSWLIIVFLTLCCGGGLGVFAFPVDTNLAIASLSFLALIYIF 540
 Db 481 EDEMTKTFAGQOETPLFQFQSWLIIVFLTLCCGGGLGVFAFPVDTNLAIASLSFLALIYIF 540
 QY 541 FIVYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPPPELQETTKALPVRFLLF 600
 Db 541 FIVYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPPPELQETTKALPVRFLLF 600
 QY 601 TDNRLSSVGGETSLSAEMIATLSDACERBFGFLATRLFRVFRTEESQGGKKWKTCCLPS 660
 Db 601 TDNRLSSVGGETSLSAEMIATLSDACERBFGFLATRLFRVFRTEESQGGKKWKTCCLPS 660
 QY 661 FVIFLFIIVGCIITLAILAIFRVDPKHLTVNAILISIASVVGFLAVLNCRNTWQVLDL 720

Db	661	VIFLFIICIGITLLAIFRVDPKHLTVNAVLISIASVVGAFVLCNRTWQVDSLL	720
Qy	721	NSQRKRLHSAKHLKSEGFMVLCKEVELMARMAKTIDSFTQONQRLVVIDGLDAC	780
Db	721	NSQRKRLHNAKHLKSEGFMVLCKEVELMARMAKTIDSFTQONQRLVVIDGLDAC	780
Qy	781	EQDKVQLMDTVRVLFSGPPIAFIPASDPHIIKAINQNLSNVLSDNSINGHDYMRNVH	840
Db	781	EQDKVQLMDTVRVLFSGPPIAFIPASDPHIIKAINQNLSNVLSDNSINGHDYMRNVH	840
Qy	841	LPVFLNSGLSNARKFLYTSATNGDITCSDTTGQEDTDRRVSONSLGEMTKLGSKTALN	900
Db	841	LPVFLNSGLSNARKFLYTSATNGDITCSDTTGQEDTDRRVSONSLGEMTKLGSKTALN	900
Qy	901	RRDYRRRQOMORTITROMSFDLTKLLVTEDFWSDISPOTRMRLLNIVSVTGRLLRANOIT	960
Db	901	RRDYRRRQOMORTITROMSFDLTKLLVTEDFWSDISPOTRMRLLNIVSVTGRLLRANOIT	960
Qy	961	FNWDLASWINLTCQWYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL	1020
Db	961	FNWDLASWINLTCQWYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL	1020
Qy	1021	EIDGDIRNFVFLSRTVPVLVARDVKTFPLCTVNLDPKRLREIADVRAAREQINIGGLAY	1080
Db	1021	EIDGDIRNFVFLSRTVPVLVARDVKTFPLCTVNLDPKRLREIADVRAAREQINIGGLAY	1080
Qy	1081	PPLPLHEGPPPPSGYSPASVCSASFNPGGVSPPOHSSVYSGLSGQPPFYNRA	1140
Db	1081	PPLPLHEGPPPPSGYSPASVCSASFNPGGVSPPOHSSVYSGLSGQPPFYNRA	1140
Qy	1141	AVPATGSSLLSMTVDVCEKLRQIEGLDQNMPOYCTTIKANINGRVLSCQNDIBLK	1200
Db	1141	SGPAGPVVLLNSLNDVACEKLRQIEGLDQNMPOYCTTIKANINGRVLSCQNDIBLK	1200
Qy	1201	KEMANFGDWHLFRSMVLEMSVSQVVPDPRFLNENSSAPVPHGESARRSHTPLT	1260
Db	1201	KEMANFGDWHLFRSMVLEMSVSQVVPDPRFLNENSSAPVPHGESARRSHTPLT	1260
Qy	1261	ELSSQTPYTLNFSFEELNLGLDEGAPRHSNLWSQOTRTPPSLSLNSQDSSIEISKLT	1320
Db	1261	ELSSQTPYTLNFSFEELNLGLDEGAPRHSNLWSQOTRTPPSLSLNSQDSSIEISKLT	1320
Qy	1321	DKVQAEYDAYREYIAQMSQLEGSGSTIGRSPSHSTYIGOSSGSGSTHSTLEQERG	1380
Db	1321	DKVQAEYDAYREYIAQMSQLEGSGSTIGRSPSHSTYIGOSSGSGSTHSTLEQERG	1380
Qy	1381	KEGELKQEDGRKSFIMKRGDVIDYSSGVSVTNEASPLDPIITEDEKSDQSGSKLLPGKKS	1440
Db	1381	KDSEPKPDGRKSFIMKRGDVIDYSSGVSVTNEASPLDPIITEDEKSDQSGSKLLPGKKS	1440
Qy	1441	SERPSLFQTDKLKGGGLRYOKLSDDESGTGRVOITPHCSKMRITKRLKAKORECASP	1500
Db	1441	SERPSLFQTDKLKGGGLRYOKLSDDESGTGRVOITPHCSKMRITKRLKAKORECASP	1500
Qy	1501	QEHSAEPIRTFIKAKYLSLALLDKQSDSDGVSRSNPSHSLHNEAADSDQLEKANLI	1560
Db	1501	PEHSAEPIRTFIKAKYLSLALLDKQSDSDGVSRSNPSHSLHNEAADSDQLEKANLI	1560
Qy	1561	ELEDEHSGKRGMPHSLSGLODPIIARMSICSEKSPSECSLIASSPEESWPACQAYN	1620
Db	1561	ELEDEHSGKRGMPHSLSGLODPIIARMSICSEKSPSECSLIASSPEESWPACQAYN	1620
Qy	1621	LNRTPTVTLLNNAPTNRANQNFDEHIGRETQVILRPGPSPNPPTAVQENILKSMHAK	1680
Db	1621	LNRTPTVTLLNNAPTNRANQNFDEHIGRETQVILRPGPSPNPPTAVQENILKSMHAK	1680
Qy	1681	RSQSSSYRLSKDASELH-AASSSTGFGGERESIL 1715	
Db	1681	RSQSSSYRLSKDASELH-AASSSTGFGGERESIL 1715	

AAM38993	
ID	AAM38993 standard; protein; 1715 AA.
XX	
AC	AAM38993;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2138.
XX	
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US034263.
XX	
PR	23-DEC-1999; 99US-00471275.
XX	
PR	21-JAN-2000; 2000US-00488725.
XX	
PR	25-APR-2000; 2000US-00552317.
XX	
PR	20-JUN-2000; 2000US-00598042.
XX	
PR	19-JUL-2000; 2000US-00620312.
XX	
PR	03-AUG-2000; 2000US-00653450.
XX	
PR	14-SEP-2000; 2000US-00662191.
XX	
PR	19-OCT-2000; 2000US-00693036.
XX	
PR	29-NOV-2000; 2000US-00727344.
XX	
FA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX	
PI	Zhou P, Goodrich R, Drmanac RT;
XX	
XX	WPI; 2001-442253/47.
DR	N-PSDB; AAI58149.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
XX	
PS	Example 4; SEQ ID NO 2138; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification
XX	
SQ	Sequence 1715 AA;

Query Match	93.4%;	Score 8266;	DB 4;	Length 1715;
Best Local Similarity	92.7%;	Pred. No. 0;		
Matches 1591;	Conservative	53;	Mismatches 70;	Indels 2; Gaps 2;
Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERCEGQTPLMLAABQGNVVEIVKELKN	60	
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERCEGQTPLMLAABQGNVVEIVKELKN	60	

QY 61 GANCNLELDLNTWALTISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120
DB 61 GANCNLELDLNTWALTISASKEGHVHIVEELLKCGVNLLEHRDMGWTALMWACYKGRDVV 120
QY 121 ELLLSHGANSVTLQVSVYPIIWAAGRGHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
DB 121 ELLLSHGANSVTLQVSVYPIIWAAGRGHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
QY 181 KGHLECVKHLANGADVQSGANSMTALIVAVKGGYTSQSVKEILKRNPNVNLTKDGNNTA 240
DB 181 KGHLECVKHLANGADVQSGANSMTALIVAVKGGYTSQSVKEILKRNPNVNLTKDGNNTA 240
QY 241 LMTASKEGHEIEIVODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
DB 241 LMTASKEGHEIEIVODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
QY 301 QDNKTALYWAVEKGNATWVDILQCNPDTEICTKOGETPLIKATKRNIEVWELLDDKGA 360
DB 301 QDNKTALYWAVEKGNATWVDILQCNPDTEICTKOGETPLIKATKRNIEVWELLDDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGSRRLAEILLRNPKDGRLLYRPNKAGETPPYINDCSHQKSL 420
DB 361 KVSADVKKGDTPLHVAIRGSRRLAEILLRNPKDGRLLYRPNKAGETPPYINDCSHQKSL 420
QY 421 TQIFGARHLSPTEITDGMGLGYDYSSALADILSEPTMOPPICVGLYAQWGSFGKSLKKL 480
DB 421 TQIFGARHLSPTEITDGMGLGYDYSSALADILSEPTMOPPICVGLYAQWGSFGKSLKKL 480
QY 481 EDEMKTFAGOOTPLFOFWSMLIVFLTLCCGLGLVFAFPVDVNTLATAISLSFLALIYIF 540
DB 481 EDEMKTFAGOOIEPLFOFWSMLIVFLTLCCGLGLLFAFTVHNLGIAVSLFLALIYIF 540
QY 541 FIVYFGRRREGESWNWAWLSRLARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLF 600
DB 541 FIVYFGRRREGESWNWAWLSRLARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLF 600
QY 601 TDYNRLLSVGGETSLAEMIATLSDACEREGFLATRLFRVFRTEESGKKKKWTKCCLPS 660
DB 601 TDYNRLLSVGGETSLAEMIATLSDACEREGFLATRLFRVFRTEESGKKKKWTKCCLPS 660
QY 661 FVIFLFTVGCIIAGITLLAIFRVPDKHLTVNALISIASVVGAFVNLNCRWTMQVLDLSL 720
DB 661 FVIFLFTIIGCIISGITLLAIFRVPDKHLTVNAVLSIASVVGAFVNLNCRWTMQVLDLSL 720
QY 721 NSQKRLHSAASKLHKLSGFGFMKVLKCEVELMARMAKTTDSFTQNTQRLVVIIDGLDAC 780
DB 721 NSQKRLHNAASKLHKLSGFGFMKVLKCEVELMARMAKTTDSFTQNTQRLVVIIDGLDAC 780
QY 781 EQDKVLQMLDTVRVLSKGPPIAIFASDPHIIKAINQNLSVLRDSNNGHDMYRNIVH 840
DB 781 EQDKVLQMLDTVRVLSKGPPIAIFASDPHIIKAINQNLSVLRDSNNGHDMYRNIVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
DB 841 LPVFLNSRGLSNARKFLVTSATNGDVPVCSDTTGTQEDADRRVSONSLGEMTKLGSKTALN 900
QY 901 RRDYRRRQMQRTITRQMSFDLTKLVTEWDFSDISQPTWRRLLINIVSVTGRLLRANQIT 960
DB 901 RRDYRRRQMQRTITRQMSFDLTKLVTEWDFSDISQPTWRRLLINIVSVTGRLLRANQIS 960
QY 961 FNWDRLASWINLTQWVPYRTSWLILYLEETEGLPDQMTLKTMYRISKNIPPTTKDVEPLL 1020
DB 961 FNWDRLASWINLTQWVPYRTSWLILYLEETEGIPDQMTLKTMYRISKNIPPTTKDVEPLL 1020
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLPCTVNLNLPKLEIITADVRAAREQINIGGLAY 1080
DB 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLNLPKLEIITADVRAAREQISIGGLAY 1080
QY 1081 PPLPLHGGPRRPSGYQPSASVCSASFNGPFGGVVSPQPHSSYSGLSGPOHPFYNRA 1140
DB 1081 PPLPLHGGPRRPSGYQPSVCSSTSFNGPFGAGVVSPQPHSSYSGMTGPQHPFYNRG 1140

QY 1141 AVPATGSSILLSSMTVDVVECKLRQIEGLDQNNMPOYCTTIKKANINGRVLSCNIDELK 1200
DB 1141 SGPAQPVVLLSLNVDVAVECKLRQIEGLDQMLPOYCTTIKKANINGRVLACNIDELK 1200
QY 1201 KEMANFQDWHLFRSVMLEMRSVESQVVPEDPRFLNENSAFVPHGESARRSHTLPLT 1260
DB 1201 KEMANFQDWHLFRSVMLEMRNAESHVVPEDPRFLSESSGPPHAPGEPARRASHNELPHT 1260
QY 1261 ELSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSTQTRTPSLSLNSQSSSIEISKLT 1320
DB 1261 ELSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSTQTRTPSLSLNSQSSSIEISKLT 1320
QY 1321 DKVQAYRDAYRYIIAQMSOLEGGTSSITSGRSSPHSTVYIQSSSGSGSIHSTLEOERG 1380
DB 1321 DKVQAYRDAYRYIIAQMSOLEGGPGSTTISGRSSPHSTVYIMQSSSGSIHNLQOEGK 1380
QY 1381 KEGELKQBDGRKSFMLMKRGDVIDYSSGVSTNEASPLDPTITERDEKSDQSGSKLLPGKKS 1440
DB 1381 KQSEPKPDGRKSFMLMKRGDVIDYSSGVSTNDASPLDPTITERDEKSDQSGSKLLPGKKS 1440
QY 1441 SERPSLFOTDLKLKGGGLRYOKLPSDEDSGTGRVQITPHCSKMIRTKLKAQORECASP 1500
DB 1441 SERSSLFOTDLKLKGGGLRYOKLPSDEDSGTGTEESDNTPLKDDKDKRKAEGKVERVPKS 1499
QY 1501 QEHSAPRTFTFIKAKELYSDALLDKKSDSDSGVRSNESPNSHLSHNEAADDQLEKANLI 1560
DB 1500 PHSASPIRTFTFIKAKELYSDALLDKKSDSDSGVRSSESPNSHLSHNEVADDQLEKANLI 1559
QY 1561 ELEDEHSGKRGMPHLSGLQDPIIARMGICSEDKSPSECSLIASSPEESWPACQKAYN 1620
DB 1560 ELEDDSHSGKRGIPHLSGLQDPIIARMGICSEDKSPSECSLIASSPENWPACQKAYN 1619
QY 1621 LNRTPTVTILNNTATNEANQNFDEIGIRETSQVILRPGPSPNTAVQENLNKSMVHK 1680
DB 1620 LNRTPTVTILNNSAPANEANQNFDEIGIRETSQVILRPPSSNPPTTIQNENLKSMTWK 1679
QY 1681 RSQRSSYTRLSDASELH-AASSESTGFGEERESIL 1715
DB 1680 RSQRSSYTRLSDKPPPELHAAASSESTGFGEERESIL 1715
RESULT 8
AAE25144
ID AAE25144 standard; protein; 1715 AA.
XX AC AAE25144;
XX DT 30-OCT-2002 (first entry)
XX DE Human ARMS protein.
XX KW Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW diagnostic; imaging; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FH Region 17..390
FT /note= "Ankyrin repeat"
FT Region 354..493
FT /note= "N-terminal region"
FT Domain 496..518
FT /note= "Transmembrane domain"
FT Domain 525..546
FT /note= "Transmembrane domain"
FT Region 573..638
FT /note= "Loop between TM2 and TM3"
FT Domain 661..680
FT /note= "Transmembrane domain"
FT Domain 688..710
FT /note= "Transmembrane domain"
FT Region 749..854

FT Region /note= "Carboxy tail"
 FT 940..1060
 FT /note= "Carboxy tail"
 FT 1081..1093
 FT /note= "Polyproline stretch"
 FT 1151..1221
 FT /note= "SAM domain"
 FT 1713..1715
 FT /note= "PDZ binding motif"
 XX
 PN WO200250273-A2.
 XX
 XX 27-JUN-2002.
 PD
 XX
 XX
 PF 19-DEC-2001; 2001WO-US048603.
 XX
 PR 21-DEC-2000; 2000US-0256909P.
 XX
 PA (UNYNY) UNIV NEW YORK STATE.
 XX
 XX Chao MV, Kong H.
 FI
 XX
 DR WPI; 2002-508800/54.
 DR N-PSDB; AAD41036.
 XX
 PT Isolated ankyrin repeat-rich membrane spanning (ARMS) polypeptide that is
 PT a target for phosphorylation by neurotrophin and ephrin receptor tyrosine
 PT kinases, useful as a marker for growth cones.
 XX
 PS Claim 1; Page 103-110; 136pp; English.
 XX
 CC The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
 CC protein which is a down stream target of neurotrophin and ephrin receptor
 CC tyrosine kinases. ARMS DNA is useful for visualising the growth cone of
 CC neurons. ARMS protein is useful as an indicator of the biological
 CC activity of neurotrophins and ephrins, as a marker for neuronal cells
 CC which have the ability to undergo continued synaptic changes through
 CC adult life or for the presence and distribution of ARMS in such neuronal
 CC cells, or as a marker for growth cones. ARMS DNA is useful in diagnostic
 CC and imaging methods. The present sequence is human ARMS protein
 XX
 SQ Sequence 1715 AA;
 Query Match 93.4%; Score 8266; DB 5; Length 1715;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
 Qy 1 MSVLISQSVINVYVEENIPALKKLECKVDNERNECQOTPLMLAAEQGNVEIVKELKN 60
 Db 1 MSVLISQSVINVYVEENIPALKKLECKVDNERNECQOTPLMLAAEQGNLEIVKELKN 60
 Qy 61 GANCNLELDNWTALISASKEGHIHIVEBELLKSGASLEHRDMGGWTALMWACYKGRDGV 120
 Db 61 GANCNLELDNWTALISASKEGHIHIVEBELLKSGVLEHRDMGGWTALMWACYKGRDGV 120
 Qy 121 ELLLSHGANSPTVGLQYVYPIIWAAGRGHADI VHLILLONGAKVNCSDKYGTTPLVWAAR 180
 Db 121 ELLLSHGANSPTVGLQYVYPIIWAAGRGHADI VHLILLONGAKVNCSDKYGTTPLVWAAR 180
 Qy 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
 Db 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
 Qy 241 LMIASKEGHIIVQDLLDAGTYVNIPOBSGDTVLIGAVRGGHVIVRALLQYADIDIRG 300
 Db 241 LMIASKEGHIIVQDLLDAGTYVNIPOBSGDTVLIGAVRGGHVIVRALLQYADIDIRG 300
 Qy 301 QDNKTALVWYVEKGNATVRILOCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 Db 301 QDNKTALVWYVEKGNATVRILOCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 Qy 361 KVSADVKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
 Db 361 KVSADVKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420

Db 361 KVSADVKKGDTPLHVAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
 Qy 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSPTMQPPICVGLYAWGSGKSLKKL 480
 Db 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSPTMQPPICVGLYAWGSGKSLKKL 480
 Qy 481 EDEMKTFAQQOQTEPLFQFWSLIVFLTLCCGGLGVFAFPVDTNLAIAISISFLALIYIF 540
 Db 481 EDEMKTFAQQOQTEPLFQFWSLIVFLTLCCGGLGVFAFPVDTNLAIAISISFLALIYIF 540
 Qy 541 FIVYFGRREGESNNWAWALSTRLARHIGYLELLFKLMFVNPPPELPTQTKALPVRLFL 600
 Db 541 FIVYFGRREGESNNWAWALSTRLARHIGYLELLFKLMFVNPPPELPTQTKALPVRLFL 600
 Qy 601 TDYNRLSSVGGETSLAEMIATLSDACREFFGLATRLFRVFTERTESQCKKWKTCCLPS 660
 Db 601 TDYNRLSSVGGETSLAEMIATLSDACREFFGLATRLFRVFTERTESQCKKWKTCCLPS 660
 Qy 661 FVIFLFIIVGCIAGITTLAIFRVDPKHLTVNAILISIASVVGGLAFVLNCRTWQVLDLIL 720
 Db 661 FVIFLFIIVGCIAGITTLAIFRVDPKHLTVNAILISIASVVGGLAFVLNCRTWQVLDLIL 720
 Qy 721 NSQKRLHSAASKLHLKSEGFVKLVCEVELMARMAMAKTIDSFQNTQRLVVIDGLDAC 780
 Db 721 NSQKRLHSAASKLHLKSEGFVKLVCEVELMARMAMAKTIDSFQNTQRLVVIDGLDAC 780
 Qy 781 BQDKVQLMDTVRVLFPSKGPPIAFASDPHPIIIKAINQNLNSVLKRDNSINGHDYMRNVH 840
 Db 781 BQDKVQLMDTVRVLFPSKGPPIAFASDPHPIIIKAINQNLNSVLKRDNSINGHDYMRNVH 840
 Qy 841 LPVFLNRSGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
 Db 841 LPVFLNRSGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
 Qy 901 RRDYRRRQMQRTITRQMSFDLTKLVTEWDFSDISPTQMRRLNINIVSVTGRLLRANQIT 960
 Db 901 RRDYRRRQMQRTITRQMSFDLTKLVTEWDFSDISPTQMRRLNINIVSVTGRLLRANQIT 960
 Qy 961 FNWDLASWNLTEQWYRYSWLLIYLFEEGLEPDMQTLKMYERISKNIPTTKDVEPLL 1020
 Db 961 FNWDLASWNLTEQWYRYSWLLIYLFEEGLEPDMQTLKMYERISKNIPTTKDVEPLL 1020
 Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTPCTVNLDPKLEIADVRAAREQINIGGLAY 1080
 Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKTPCTVNLDPKLEIADVRAAREQINIGGLAY 1080
 Qy 1081 PPLPLHEGPPRPSPGYSQPAVCSASFNPGFPVGVSPQPHSSYVYSGLSGQPHFPYNRA 1140
 Db 1081 PPLPLHEGPPRPSPGYSQPAVCSASFNPGFPVGVSPQPHSSYVYSGLSGQPHFPYNRA 1140
 Qy 1141 AVPATGSSLLLSMTVDVCEKLEKLEGLDQNMMPQVCTTIKKANINGRVLSCNIDELK 1200
 Db 1141 AVPATGSSLLLSMTVDVCEKLEKLEGLDQNMMPQVCTTIKKANINGRVLSCNIDELK 1200
 Qy 1201 KEMANFGDWHLPFSWYLEMRSVESQVVPEDPRFLNENSAAPVPHGSARSSTHTEPLT 1260
 Db 1201 KEMANFGDWHLPFSWYLEMRSVESQVVPEDPRFLNENSAAPVPHGSARSSTHTEPLT 1260
 Qy 1261 ELSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRTPSLSSLSNSQDSSIEISKLT 1320
 Db 1261 ELSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRTPSLSSLSNSQDSSIEISKLT 1320
 Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYIIGOSSGSGSHSTLEQBERG 1380
 Db 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYIIGOSSGSGSHSTLEQBERG 1380
 Qy 1381 KEGSLKQEDGRKSPMKRGDVIDYSSSGVSTNEASPLDPTITERDEKSDQSGKLLPGKKS 1440
 Db 1381 KEGSLKQEDGRKSPMKRGDVIDYSSSGVSTNEASPLDPTITERDEKSDQSGKLLPGKKS 1440
 Qy 1441 SERPSLFQTDLKLKGGGLRYQKLPSPDESGTGVRQITPHCSKMIRTKRLKAKQREKASP 1500
 Db 1441 SERPSLFQTDLKLKGGGLRYQKLPSPDESGTGVRQITPHCSKMIRTKRLKAKQREKASP 1500

QY 781 EODKVLQMLDTRVRLFSKGPFIAPSDPHIIKAINQNLSVLRSDNSINGHDYMRNIWH 840
 Db 781 EODKVLQMLDTRVRLFSKGPFIAPSDPHIIKAINQNLSVLRSDNSINGHDYMRNIWH 840
 QY 841 LPVFLNSGLSNARKFLVTSATNGDITCSDDTGTQEDTDRVSQNSLGMTKLSKLTALN 900
 Db 841 LPVFLNSGLSNARKFLVTSATNGDITCSDDTGTQEDTDRVSQNSLGMTKLSKLTALN 900
 QY 901 RRDYRRRQMQRTITROMSFOLTKLLVTFEDFSDISPTMRRLNIVSVTGRLLRANOIT 960
 Db 901 RRDYRRRQMQRTITROMSFOLTKLLVTFEDFSDISPTMRRLNIVSVTGRLLRANOIS 960
 QY 961 FNVDRLASWINLQWVPYRTSWLILYLETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 Db 961 FNVDRLASWINLQWVPYRTSWLILYLETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 QY 1021 EIDGDIRNFVFLSRTPVLVARDVKFPLCTVNLDPKRLIADVRAAREQINIGGLAY 1080
 Db 1021 EIDGDIRNFVFLSRTPVLVARDVKFPLCTVNLDPKRLIADVRAAREQISIGGLAY 1080
 QY 1081 PPLPLHEGPPPPSGYSPASVCSASFNGPFGVSPQPHSSVYSGLSGPQHPFYNRA 1140
 Db 1081 PPLPLHEGPPPPSGYSPASVCSASFNGPFGVSPQPHSSVYSGLSGPQHPFYNRG 1140
 QY 1141 AVPATGSLLSMTVDVVCCKLRQIEGLDQMMPOYCTTIKKANINGRVLSCQNIDELK 1200
 Db 1141 SGAPAGPVVLLNSLNDVACEKLRQIEGLDQMLPOYCTTIKKANINGRVLACQNIDELK 1200
 QY 1201 KEMANFGDWHLFRSVLMEKRSVSVQVPPDFRFLNENSSAPVPHGESARRSHTPLT 1260
 Db 1201 KEMANFGDWHLFRSVLMEKRSVSVQVPPDFRFLNENSSAPVPHGESARRSHTPLT 1260
 QY 1261 ELSSQTPYTLNFSPEELNTGLDGAPRHSNLWSQOTRRTPSLSSLSQDSSSIEISKLT 1320
 Db 1261 ELSSQTPYTLNFSPEELNTGLDGAPRHSNLWSQOTRRTPSLSSLSQDSSSIEISKLT 1320
 QY 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGSIHSTLEQERG 1380
 Db 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGSIHSTLEQERG 1380
 QY 1381 KEGELKQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
 Db 1381 KDSBPKPDDGRKSFMLKRGDVIDYSSGVSTNDASPLDPTTEDEKSDQSGSKLLPGKKS 1440
 QY 1441 SERPSLFTDLKLGKGLRYOKLPSEDESGTGRVQITPHCSKMRITKELKAKQREKASP 1500
 Db 1441 SERPSLFTDLKLGKGLRYOKLPSEDESGTGRVQITPHCSKMRITKELKAKQREKASP 1500
 QY 1501 QEHSAPERTPIKAKEVLSDALDKDSSDGVSRNNESSPHSLHNEAADDQLEKANLI 1560
 Db 1500 PEHSAPERTPIKAKEVLSDALDKDSSDGVSRNNESSPHSLHNEAADDQLEKANLI 1559
 QY 1561 ELEDGHSKRGKPHSLSGLODPIIARMSICSEKSPSECSLIASSPESWPACQAYN 1620
 Db 1560 ELEDGHSKRGKPHSLSGLODPIIARMSICSEKSPSECSLIASSPESWPACQAYN 1619
 QY 1621 LNRTPSTVTLNNAPTNRANQNFDEIEGIRETSQVILRPGPSNPATVQENILKSAHK 1680
 Db 1620 LNRTPSTVTLNNAPTNRANQNFDEIEGIRETSQVILRPSSSPNPTTIQENILKSMTHK 1679
 QY 1681 RSQSSVTRLSKDASELH-AASSESTGFERESIL 1715
 Db 1680 RSQSSVTRLSKDPPELHAASSESTGFERESIL 1715

RESULT 10

AAE32128

ID AAE32128 standard; protein; 1715 AA.

XX AAE32128;

XX 24-MAR-2003 (first entry)

XX DE Human cytoskeleton-associated protein, CSAP-26.
 XX KW Human; cytoskeleton-associated protein; CSAP-26; atherosclerosis; cancer;
 XX OS gene therapy.
 XX FH Homo sapiens.
 XX Key Location/Qualifiers
 FT Domain 42..96
 FT Domain /note= "Domain present in ZO-1"
 FT Domain 354..392
 FT Domain /note= "Domain present in ZO-1"
 FT Binding-site 467..474
 FT Domain /note= "ATP/GTP-binding site motif A"
 FT Domain 494..514
 FT Domain /note= "Transmembrane domain"
 FT Domain 524..544
 FT Domain /note= "Transmembrane domain"
 FT Domain 654..674
 FT Domain /note= "Transmembrane domain"
 FT Domain 687..707
 FT Domain /note= "Transmembrane domain"
 XX WO200279404-A2.
 XX 10-OCT-2002.
 XX 25-MAR-2002; 2002WO-US009288.
 XX 29-MAR-2001; 2001US-0280508P.
 XX 03-APR-2001; 2001US-0281323P.
 XX 13-APR-2001; 2001US-0283769P.
 XX 04-MAY-2001; 2001US-0288609P.
 XX 10-MAY-2001; 2001US-0290518P.
 XX 18-MAY-2001; 2001US-0291870P.
 XX 29-MAY-2001; 2001US-0294451P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Hafalia AJA, Tang TY, Yue H, Khan FA, Ison CH, Baughn MR;
 XX Warren BA, Duggan BM, Thangavelu K, Honchell CD, Azimzai Y;
 XX Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
 XX Richardson TW, Lee SY, Bandman O, Lal PG, Lee S, Gietzen KJ;
 XX Walia NK, Griffin JA, Lee EA, Swarnakar A, Ring HZ, Jones KA;
 XX WPI; 2003-092894/08.
 XX N-PSDB; AAD49615.
 XX New human cytoskeleton-associated proteins, useful for preparing a
 XX composition for diagnosing or treating a disease or condition associated
 XX with decreased expression or overexpression of functional CSAP e.g.,
 XX cancer.
 XX Claim 1; Page 195-199; 233pp; English.
 XX The invention relates to new human cytoskeleton-associated protein (CSAP)
 XX and its polynucleotide. The polypeptide is useful for preparing a
 XX composition for diagnosing or treating a disease or condition associated
 XX with decreased expression or overexpression of functional CSAP e.g.
 XX atherosclerosis or cancer. The present sequence is human CSAP-26 protein.
 XX The invention is useful in gene therapy
 XX Sequence 1715 AA;
 XX Query Match 93.4%; Score 8266; DB 6; Length 1715;
 XX Best Local Similarity 92.7%; Pred. No. 0;
 XX Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
 QY 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN 60
 Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN 60

QY 61 GANCNLELDNWTALISASKEGHIHIVEELLKSGASLEHRDMGGTALMWAACYKGRDYY 120
DB 61 GANCNLELDNWTALISASKEGHIHIVEELLKCGVNLHRDMGGTALMWAACYKGRDYY 120
QY 121 ELLLSHGANSVTLGQVSVYPIIWAAGRHADIVHLLQNGAKVNSDKYGTTPPLVWAAR 180
DB 121 ELLLSHGANSVTLGQVSVYPIIWAAGRHADIVHLLQNGAKVNSDKYGTTPPLVWAAR 180
QY 181 KGHLECVKHLANGADVQEGANSMTALIVAVKGGYTQSVKELTKRNPVNLTKDGN 240
DB 181 KGHLECVKHLANGADVQEGANSMTALIVAVKGGYTQSVKELTKRNPVNLTKDGN 240
QY 241 LMTASKEGHIHIVEELLKSGASLEHRDMGGTALMWAACYKGRDYY 120
DB 241 LMTASKEGHIHIVEELLKSGASLEHRDMGGTALMWAACYKGRDYY 120
QY 301 QDNKNTALYWAVERKGNATMVRDILQCNPDTEICTKGGETPLIKATKVRNIEVVELLDKGA 360
DB 301 QDNKNTALYWAVERKGNATMVRDILQCNPDTEICTKGGETPLIKATKVRNIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLNPKDGRLLYRPNKAGETPNIDCSHOKSIL 420
DB 361 KVSADVKKGDTPLHVAIRGSRRLAELLNPKDGRLLYRPNKAGETPNIDCSHOKSIL 420
QY 421 TQIFGARHLSPTETDGMGLDYLSALADILSEPTWQPPICVGLYAQWGSKGSFLK 480
DB 421 TQIFGARHLSPTETDGMGLDYLSALADILSEPTWQPPICVGLYAQWGSKGSFLK 480
QY 481 EDEMKTFAGQOIEPLFQFWSLIVFLTLCCGGLLFAFTVHNLGIAVSLSPALLIYIF 540
DB 481 EDEMKTFAGQOIEPLFQFWSLIVFLTLCCGGLLFAFTVHNLGIAVSLSPALLIYIF 540
QY 541 FIVIVYGGRRGEGSWNVAWLSFLARHIGYLELLKFMFVNPPELPEQTTKALPVRF 600
DB 541 FIVIVYGGRRGEGSWNVAWLSFLARHIGYLELLKFMFVNPPELPEQTTKALPVRF 600
QY 601 TDYNRSLSSVGGTSLAEMIATLSDACERBFGFLATLFRVFRTEESQKKWKTKCCLPS 660
DB 601 TDYNRSLSSVGGTSLAEMIATLSDACERBFGFLATLFRVFRTEESQKKWKTKCCLPS 660
QY 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGFLAVLNCRTWQVLDL 720
DB 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGFLAVLNCRTWQVLDL 720
QY 721 NSQKRLHNSAASKHLKSEGFNKLKCEVELMARMAKTIDSTQNTQRLVWIIDGLD 780
DB 721 NSQKRLHNSAASKHLKSEGFNKLKCEVELMARMAKTIDSTQNTQRLVWIIDGLD 780
QY 781 EODKVLQMLDITVRVLSKGFPIAFASDPHIIKAINQNLNSVLRDSNNGHDMYRNIVH 840
DB 781 EODKVLQMLDITVRVLSKGFPIAFASDPHIIKAINQNLNSVLRDSNNGHDMYRNIVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTQEDTDRRVQNSGLGEMTKLGSKTALN 900
DB 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTQEDTDRRVQNSGLGEMTKLGSKTALN 900
QY 901 RRDYTRRRQMQRTITQMSFDLTKLIVTEDWFDSDISPTQNRLLNIVSVTGRLLRANQIT 960
DB 901 RRDYTRRRQMQRTITQMSFDLTKLIVTEDWFDSDISPTQNRLLNIVSVTGRLLRANQIT 960
QY 961 FNDRLASWINLQEQPYRTSWILVLEETEGLPDQWTLKMYERISKNIPTTKDVEPLL 1020
DB 961 FNDRLASWINLQEQPYRTSWILVLEETEGLPDQWTLKMYERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIRNFEVFLSSRTPVLVARDVKTFLPCTVNLDPKLEIITADVRAAREQINICGLAY 1080
DB 1021 EIDGDIRNFEVFLSSRTPVLVARDVKTFLPCTVNLDPKLEIITADVRAAREQINICGLAY 1080
QY 1081 PPLPLHGGPPRPSGVSQPASVCSASFNGPFGGVVSPQPHSSYSGISGPOHPFNRA 1140
DB 1081 PPLPLHGGPPRPSGVSQPASVCSASFNGPFGGVVSPQPHSSYSGISGPOHPFNRA 1140
QY 1141 AVPATGSSLLSSMTVDVCEKRLQIEGLDQNMMPQVCTTIKKANINGRVLSCQNDL 1200

DB 1141 SGPAQPVVLLNSLVDAVCEKRLQIEGLDQNMMPQVCTTIKKANINGRVLSCQNDL 1200
QY 1201 KEMANNFGDWHLFRSMVLEMRSVESQVVPEDPRPLNENSSAPVPHGESARRSHTLPLT 1260
DB 1201 KEMANNFGDWHLFRSMVLEMRSVESQVVPEDPRPLNENSSAPVPHGESARRSHTLPLT 1260
QY 1261 ELSSQTPYTLNFSPELNTLGLDEGAPRHSNLSWOSQTRRTPSLSLSNQDSSIEISKLT 1320
DB 1261 ELSSQTPYTLNFSPELNTLGLDEGAPRHSNLSWOSQTRRTPSLSLSNQDSSIEISKLT 1320
QY 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSPHSTYIICQSSSGGSIHSTLEQERG 1380
DB 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSPHSTYIICQSSSGGSIHSTLEQERG 1380
QY 1381 KGEGLKQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDITEDEKSDQSGSKLLPGKKS 1440
DB 1381 KGEGLKQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDITEDEKSDQSGSKLLPGKKS 1440
QY 1441 SERPSLFOYDLKXGGLRYQKLPSEDESGTGRVQITPHCSKMIKRTKELKAKQRECAP 1500
DB 1441 SERPSLFOYDLKXGGLRYQKLPSEDESGTGRVQITPHCSKMIKRTKELKAKQRECAP 1500
QY 1501 QHSAEPIRTFIKAEYLSKDALDKKSDSGVRSNENSSPHSLHNEAADDQSOLEKANLI 1560
DB 1501 QHSAEPIRTFIKAEYLSKDALDKKSDSGVRSNENSSPHSLHNEAADDQSOLEKANLI 1560
QY 1561 ELEDGHSKGRGMPHSLGLOPPIIARMSICEDKKSPECSLIASSPESWPACOKAYN 1620
DB 1561 ELEDGHSKGRGMPHSLGLOPPIIARMSICEDKKSPECSLIASSPESWPACOKAYN 1620
QY 1621 LNRSTSTVTLNNTAPTNRANONFDEIGIRTSOVLIRPSPNPPTAVONENLKSMWK 1680
DB 1621 LNRSTSTVTLNNTAPTNRANONFDEIGIRTSOVLIRPSPNPPTAVONENLKSMWK 1680
QY 1681 RSQRSSYTLKSDASELH-AASSESTGFGEERESIL 1715
DB 1681 RSQRSSYTLKSDASELH-AASSESTGFGEERESIL 1715
RESULT 11
ADA09888
ID ADA09888 standard; protein; 1753 AA.
XX AC ADA09888;
XX DT 20-NOV-2003 (first entry)
XX DE Human receptor and membrane-associated protein REMAP-30, SEQ ID:30.
XX KW Human; receptor and membrane-associated protein; REMAP;
XX KW cell proliferative disorder; cancer; autoimmune disorder;
XX KW inflammatory disorder; infection; neurological disorder;
XX KW metabolic disorder; developmental disorder; endocrine disorder;
XX KW cytostatic; immunosuppressive; antiinflammatory; neuroprotective;
XX KW nootropic; cerebroprotective; gene therapy; receptor.
OS Homo sapiens.
PN WO2003070902-A2.
XX PD 28-AUG-2003.
XX PF 18-FEB-2003; 2003WO-US004902.
XX PR 20-FEB-2002; 2002US-0358279P.
XX PR 13-MAR-2002; 2002US-0364338P.
XX PR 25-APR-2002; 2002US-0375657P.
XX PR 29-APR-2002; 2002US-0376669P.
XX PR 10-MAY-2002; 2002US-0379837P.
XX PR 10-MAY-2002; 2002US-0379853P.
XX PA (INCY-) INCYTE GENOMICS INC.

XX Chawla NK, Yue H, Richardson TW, Marquis JP, Lehr-Mason PM;
 PI Gorvad AE, Becha SD, Kable AE, Swarnakar A, Jin P, Hawkins PR;
 PI Chien D, Ramkumar J, Tran UK, Hafalia AJA, Baughn MR, Lee SY;
 PI Jiang X, Jackson AA, Khare R, Bulloch SA;
 XX WPI; 2003-697610/66.
 DR N-PSDB; ADA09926.
 XX
 PT New human receptor and membrane associated proteins and nucleic acids,
 PT useful for diagnosing, treating or preventing e.g. viral, bacterial,
 PT fungal, parasitic, protozoan or helminthic infections, cancers,
 PT neurological disorders.
 XX
 PS Claim 1; Page 254-257; 298pp; English.
 XX
 CC The invention relates to 38 human receptors and membrane-associated
 CC proteins (REMAP), REMAP-1 to REMAP-38, and the cDNAs encoding them
 CC (ADA09859-ADA09934). The invention also encompasses expression
 CC constructs, host cells and transgenic organisms comprising a REMAP
 CC nucleic acid sequence; the recombinant preparation of a REMAP; an
 CC antibody against a REMAP; methods of detection of REMAP proteins or
 CC nucleic acids; a micro-array containing REMAP nucleic acids; methods of
 CC screening compounds for their ability to modulate REMAP activity or
 CC expression; and pharmaceutical compositions comprising a REMAP protein, a
 CC REMAP agonist or REMAP antagonist. The REMAP proteins, nucleic acids or
 CC compositions comprising them are useful in diagnosing, treating or
 CC preventing a variety of disorders, including cell proliferative disorders
 CC (e.g., arteriosclerosis, cirrhosis, hepatitis, psoriasis, or primary
 CC thrombocytopenia) or cancers (e.g., adenocarcinoma, leukaemia, or cancers
 CC of the bone, brain, breast or uterus); autoimmune/inflammatory disorders
 CC (e.g., AIDS, allergies, anaemia, asthma, bronchitis, gout, multiple
 CC sclerosis, osteoarthritis, rheumatoid arthritis, or systemic lupus
 CC erythematosus); viral, bacterial, fungal, parasitic, protozoan or
 CC helminthic infections; neurological disorders (e.g., epilepsy, stroke,
 CC Alzheimer's disease, dementia, or Parkinson's disease); metabolic
 CC disorders (e.g., hereditary fructose intolerance, obesity, hypoglycaemia
 CC or diabetes); developmental disorders (e.g., achondroplastic dwarfism,
 CC hypothyroidism or hydrocephalus); or endocrine disorders (e.g., disorders
 CC of the hypothalamus or pituitary gland). The present sequence represents
 CC a REMAP of the invention.
 XX
 SQ Sequence 1753 AA;

Query Match 93.0%; Score 8237; DB 6; Length 1753;
 Best Local Similarity 90.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 40; Gaps 3;

Qy 1. MSVLISQSVINVYBEENIPALKALLECKVDNERNECQOTPLMLAAEQGNVEIVKELIKN 60
 Db 1 MSVLISQSVINVYBEENIPALKALLECKVDNERNECQOTPLMLAAEQGNVEIVKELIKN 60

Qy 61 GANCNLELDNWTALISAKGHIHIVBELLSKSGASLEHRDMGGWTALMWACYKGRITDVV 120
 Db 61 GANCNLELDNWTALISAKGHIHIVBELLSKSGVLEHRDMGGWTALMWACYKGRITDVV 120

Qy 121 ELLLSHGANSPTVGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYGTTPLVWAAR 180
 Db 121 ELLLSHGANSPTVGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYGTTPLVWAAR 180

Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTDKGNNTA 240
 Db 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTDKGNNTA 240

Qy 241 LMIAKSGHIHIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVIRVALLQYADIDIRG 300
 Db 241 LMIAKSGHIHIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVIRVALLQYADIDIRG 300

Qy 301 QDNKTALYWAKEGNATWVRDILOCNPDTEICTKDGTEPLIKATMNRNIEVVELLDKGA 360
 Db 301 QDNKTALYWAKEGNATWVRDILOCNPDTEICTKDGTEPLIKATMNRNIEVVELLDKGA 360

Qy 361 KVSADVKKGGDTPFLHVAIRGRSRLAELLRLNPKDGRLLYRPNKAGETPYNIDCSHOKSIL 420

Db 361 KVSADVKKGGDTPFLHVAIRGRSRLAELLRLNPKDGRLLYRPNKAGETPYNIDCSHOKSIL 420
 Qy 421 TQIFGARHLSPETETGDMGLYSSALADILSPTWQPPICVGLYQAQWGSGKSFLLKKL 480
 Db 421 TQIFGARHLSPETETGDMGLYSSALADILSPTWQPPICVGLYQAQWGSGKSFLLKKL 480
 Qy 481 EDEMKTFAGQOTEPLFQFWSMLIVELTLLCCGGLGVFAFPVDTNLAIASISFLALIYIF 540
 Db 481 EDEMKTFAGQOTEPLFQFWSMLIVELTLLCCGGLGVFAFPVDTNLAIASISFLALIYIF 540
 Qy 541 FIVTYFGRREGESWNWAWALSTRLARHIGYLELLFKLMFVNPPPELPTTKALPVRELF 600
 Db 541 FIVTYFGRREGESWNWAWALSTRLARHIGYLELLFKLMFVNPPPELPTTKALPVRELF 600
 Qy 601 TDYNRLSVGGETSIAEMIATLSDACERERGFATRLFRVPTESQKKKKKTKCCLPS 660
 Db 601 TDYNRLSVGGETSIAEMIATLSDACERERGFATRLFRVPTESQKKKKKTKCCLPS 660
 Qy 661 FVIFLFTVGCIIAGITILLAIFRVDPKELTNNAILSIASVVGFLAPVLCNRTWQVLDSSL 720
 Db 661 FVIFLFTVGCIIAGITILLAIFRVDPKELTNNAILSIASVVGFLAPVLCNRTWQVLDSSL 720
 Qy 721 NSQKRLHSAASKLHLKSEGMKVLKCEVELMARMARKTIDSFQNTQRLVVIIDGLDAC 780
 Db 721 NSQKRLHSAASKLHLKSEGMKVLKCEVELMARMARKTIDSFQNTQRLVVIIDGLDAC 780
 Qy 781 EQDKVLOMLDTRVRLFSKGPPIAIFASDPHIIKAIQNQLNSVLNDSNINGHDYMRNIVH 840
 Db 781 EQDKVLOMLDTRVRLFSKGPPIAIFASDPHIIKAIQNQLNSVLNDSNINGHDYMRNIVH 840
 Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLASKTALN 900
 Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLASKTALN 900
 Qy 901 RRDYTRRRQMTITRQMSFDLTLLVTEDFWFSISPTQWRRLNINIVTGRLLRANOIT 960
 Db 901 RRDYTRRRQMTITRQMSFDLTLLVTEDFWFSISPTQWRRLNINIVTGRLLRANOIT 960
 Qy 961 FNNDRLASWINLFEOWPYRTSWLLIYLBETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 Db 961 FNNDRLASWINLFEOWPYRTSWLLIYLBETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020

Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIIADVRAAREQINIGGLAY 1080
 Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIIADVRAAREQINIGGLAY 1080

Qy 1081 PPLPLHEGPPRPSPGYSQASVCSASFNGPFPQGVVSPQPHSSYYSGLSQPHFPYNR - 1139
 Db 1081 PPLPLHEGPPRPSPGYSQASVCSASFNGPFPQGVVSPQPHSSYYSGLSQPHFPYNR 1140

Qy 1140 -----AAMPATGSSLLLSMTVDVCEK 1162
 Db 1141 FFAPLYTPRYPGSQHLSRPSVKTSPLRQDNGLGSGPAPGPVVVLLSLNDAVCEK 1200

Qy 1163 LRQIEGLDQNMPOYCTTIKKANINGRVLSCNIDELKEMANMFGDWHLPFRSMVLEWS 1222
 Db 1201 LKQIEGLDQNMPOYCTTIKKANINGRVLACNIDELKEMANMFGDWHLPFRSMVLEWRN 1260

Qy 1223 VESQVWPEDPFRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTLNFSFEELNTIGL 1282
 Db 1261 AESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHELSSQTPYTLNFSFEELNTIGL 1320

Qy 1283 DEGAHRSNLSWSQTRTPSLSLNSQDSISIEISKULTKVQABYRAYREYIAQMSOLE 1342
 Db 1321 DEGAHRSNLSWSQTRTPSLSLNSQDSISIEISKULTKVQABYRAYREYIAQMSOLE 1380

Qy 1343 GGTGSSITISGRSPHSTYTYGQSSSGGSIHSTLEOERKGEKQLQEDGRKSFMLKRGDVI 1402
 Db 1381 GGTGSSITISGRSPHSTYTYGQSSSGGSIHSTLEOERKGEKQLQEDGRKSFMLKRGDVI 1440

Qy 1403 DYSSSGVSVTNEASPLDPIBDEDEKSDQSGSKLAPGKKSRRPSLFTQDLKLGGLRYQK 1462

Db 1441 DYSSGVSTNDASPLDPIETHEDEKSDQSGSKLPGKSSERSLSFQTDLLKXGSLRYOK 1500
Qy 1463 LPSEDESGTGRVOITPHCSKMITRKLKAKQREKASPOHSAEPITRTFKAKKEYLSDAL 1522
Db 1501 LPSEDESGTEESDNTF-LLKDDKDKAEKGVKVERVPKSPHSAEPITRTFKAKKEYLSDAL 1559
Qy 1523 LDKKSDSGVRNESPNSHNEAADDQLEKANIIELEDGSHGSKRGMPHSLGLOD 1582
Db 1560 LDKKSDSGVRNESPNSHNEAADDQLEKANIIELEDGSHGSKRGIPHSLSGLOD 1619
Qy 1583 PIARMSICSEDKSPSECSLIASSPESWPAQKAYNLNRTPTSTVTNNAPTNRANQ 1642
Db 1620 PIARMSICSEDKSPSECSLIASSPESWPAQKAYNLNRTPTSTVTNNAPSANRANQ 1679
Qy 1643 NFDEIGIRTSQVILRPGSPNPTAVQNLNKSMAHQRQSRSTYRLSKDSELH-AAS 1701
Db 1680 NFDEMEGIRETSQVILRPPSSPNPTTQENLNKSMTHKRSQSRSTYRLSKDPPELHAAS 1739
Qy 1702 SESTGGEERESIL 1715
Db 1740 SESTGGEERESIL 1753

RESULT 12
ADC06847
ID ADC06847 standard; protein; 1771 AA.
XX AC ADC06847;
DT 18-DEC-2003 (first entry)
XX Human Kidins220Pc protein AB033076.
XX cytosolic; prostate cancer; breast; gene therapy; transgenic; human;
KW Kidins220Pc; kinase D-interacting substrate of 220kDa; chromosome 2p25.1.
XX Homo sapiens.
XX WO2003064599-A2.
XX 07-AUG-2003.
XX 24-JAN-2003; 2003WO-US001943.
XX 25-JAN-2002; 2002US-00054935.
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00141198.
PR 19-JUN-2002; 2002US-00197824.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;
XX WPI; 2003-679495/64.

New isolated polynucleotide related to cancer genes, useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating cancers, e.g. breast and prostate cancers.

Disclosure; Fig 20; 128pp; English.

The invention relates to a novel isolated polynucleotide comprising a differentially-regulated mammalian cancer gene. The polynucleotides of the invention demonstrate cytostatic activity and are differentially expressed in prostate cancer. The polynucleotide, polypeptides and methods of the invention may be useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating cancers, particularly breast and prostate cancers. Furthermore, the invention may be utilised during gene therapy procedures or in the production of transgenic animals. The current sequence is that of the prostate cancer-related protein of the invention. The current sequence is not fully

CC legible within the specification; the complete sequence was obtained from GenBank.

XX Sequence 1771 AA;
SQ

Query Match 92.8%; Score 8212; DB 7; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;

Qy 1 MSVLSQSIVNVYVEENIPALKKALEKCKVDVERNECCGOTPLMAAEOGNVEIVKELLKN 60
Db 1 MSVLSQSIVNVYVEENIPALKKALEKCKVDVERNECCGOTPLMAAEOGNVEIVKELLKN 60
Qy 61 GANCNLEDNNTALISASKEGHIHIVEELLKSGASLEHRDGMGTALMAWACYKGRDVV 120
Db 61 GANCNLEDNNTALISASKEGHIHIVEELLKSGASLEHRDGMGTALMAWACYKGRDVV 120
Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVMAAR 180
Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVMAAR 179
Qy 181 KGLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILKKNPNVNLTKDGNTA 240
Db 180 KGLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILKKNPNVNLTKDGNTA 239
Qy 241 LMTASKEGHIETVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
Db 240 LMTASKEGHIETVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
Qy 301 QDNKTALYWAVERKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
Db 300 QDNKTALYWAVERKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPLHVAIRGSRRLABELLNRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHVAIRGSRRLABELLNRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 419
Qy 421 TQIFGARHLSPTETDGMGLDYSSALADILSEPTWQPPICVGLYAQWGSFGSKFLKKL 480
Db 420 TQIFGARHLSPTETDGMGLDYSSALADILSEPTWQPPICVGLYAQWGSFGSKFLKKL 479
Qy 481 EDEMTKTFAGQOTEPFLQFQSWLIVFLTLCCGGLGVLPFAPVDPTNLAIASLSFLALIYIF 540
Db 480 EDEMTKTFAGQOTEPFLQFQSWLIVFLTLCCGGLGVLPFAPVDPTNLAIASLSFLALIYIF 539
Qy 541 FIVIFYGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPELPPEOTTKALPVRFLF 600
Db 540 FIVIFYGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPELPPEOTTKALPVRFLF 599
Qy 601 TDYNRLSSVGGETSLEAMTATLSDACERBFGFLATRLFRVFRTEESQGGKKWKTCCPLS 660
Db 600 TDYNRLSSVGGETSLEAMTATLSDACERBFGFLATRLFRVFRTEESQGGKKWKTCCPLS 659
Qy 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTNAVLISIASVVGFLAVNLCTRWQVLDLSLL 720
Db 660 FVIFLFIIVGCIAGITLLAIFRVDPKHLTNAVLISIASVVGFLAVNLCTRWQVLDLSLL 719
Qy 721 NSQKRLHSAASKHLKSEGFPMKVLKCEVELMARMAKTIDSTONQTRLWIIDGLDAC 780
Db 720 NSQKRLHSAASKHLKSEGFPMKVLKCEVELMARMAKTIDSTONQTRLWIIDGLDAC 779
Qy 781 EQDKVLQMLDTRVRLFSKGPFIASFADPHI I I KAINQNLNSVLRDSNINNGHYDMRNIVH 840
Db 780 EQDKVLQMLDTRVRLFSKGPFIASFADPHI I I KAINQNLNSVLRDSNINNGHYDMRNIVH 839
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSTPTTQEDTDRVNSQNSLGEMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSTPTTQEDTDRVNSQNSLGEMTKLGSKTALN 899
Qy 901 RRTYRRRQMRITRQMSFDLTKLIVTEDFSDI SPQTMRLNIVSVTGRLLRANQIT 960
Db 900 RRTYRRRQMRITRQMSFDLTKLIVTEDFSDI SPQTMRLNIVSVTGRLLRANQIS 959

QY 961 FNDRLASWINTQWPYRTSLILYLETEGLPQMTLKTMYERISKNIPTTKDVEPLL 1020
Db FNDRLASWINTQWPYRTSLILYLETEGLPQMTLKTMYERISKNIPTTKDVEPLL 1019
QY 1021 EIDGDIRPEVLSRTPLVARDVKTFLPCTVNLDPKLEIADVRAAREQINTGGLAY 1080
Db EIDGDIRPEVLSRTPLVARDVKTFLPCTVNLDPKLEIADVRAAREQINTGGLAY 1079
QY 1081 PPLPLHEGPPRPPSGYSPASVCSASFNPGPGVSPQPHSSYVGLSGPQHPFYN- 1138
Db PPLPLHEGPPRPPSGYSPASVCSASFNPGPGVSPQPHSSYVGLSGPQHPFYNRP 1139
QY 1139 -----RAAVP 1143
Db FAPYLYTPRYYPGSGHLSRPSVKTSILPRDQNGLEVIEDAAGLSPTDSRGSGP 1199
QY 1144 ATGSSLLSSMTVDVCEKLRQIEGLDQNMPOYCTTIKKANINGRVLSQCNIDELKKEM 1203
Db APGVVVLNSLVNDVACBKLQIEGLDQMLPOYCTTIKKANINGRVLAQCNDIDELKKEM 1259
QY 1204 AMNFGDMLFRSMVLEMRSVESQVPPEDPRFLNENSSAPVPHGESARRSSHTELPLELS 1263
Db AMNFGDMLFRSMVLEMRSVESQVPPEDPRFLNENSSAPVPHGESARRSSHTELPLELS 1319
QY 1264 SOTPYTLNFSPEELNTGLDEGAPRHSNLWSQOTRTPSLSSLSQDSSIEISKLTDKV 1323
Db SOTPYTLNFSPEELNTGLDEGAPRHSNLWSQOTRTPSLSSLSQDSSIEISKLTDKV 1379
QY 1324 QAEYRDAYREYIAQMSQLEGGTSGTISGRSPHSTYIYGSSSGGSHSTLEQBRGEG 1383
Db QAEYRDAYREYIAQMSQLEGGTSGTISGRSPHSTYIYGSSSGGSHSTLEQBRGEG 1439
QY 1384 ELKQDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKSSER 1443
Db EPKPDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKSSER 1499
QY 1444 PSLFOTDLKLGGLRYQKLPSEDESGTGRVQIIPHCCKMIRTKLKAQREKASPOEH 1503
Db SSLFOTDLKLGGLRYQKLPSEDESGTGRVQIIPHCCKMIRTKLKAQREKASPOEH 1558
QY 1504 SAERTITIKAYELSDALLDKDSDGVSNSPNSHLSNEAADSDQLEKANLIBLE 1563
Db SAERTITIKAYELSDALLDKDSDGVSNSPNSHLSNEAADSDQLEKANLIBLE 1618
QY 1564 DEHSGKRGMPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESWPACOKAYNLNR 1623
Db DSHSGKRGMPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESWPACOKAYNLNR 1678
QY 1624 TPSTVTLNNTAPTNRANQNFDEIGIRETSQVILRPGFSPNPTAVQENILKSMHRSQ 1683
Db TPSTVTLNNTAPTNRANQNFDEIGIRETSQVILRPGFSPNPTAVQENILKSMHRSQ 1738
QY 1684 RSVYTRLSKDASELH-AASSESTGGERESIL 1715
Db RSVYTRLSKDASELH-AASSESTGGERESIL 1771

RESULT 13

ID ADC06850 standard; protein; 1142 AA.

XX AC ADC06850;

XX DT 18-DEC-2003 (first entry)

XX DE Kidins 220Pc-related protein Pc473.

XX KW cytosolic; prostate cancer; breast; gene therapy; transgenic;

XX KW Kidins220Pc; kinase D-interacting substrate of 220KDa.

XX OS Unidentified.

XX FN WO2003064599-A2.

XX 07-AUG-2003.
XX 24-JAN-2003; 2003WO-US001943.
XX 25-JAN-2002; 2002US-00054935.
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00144198.
PR 19-JUL-2002; 2002US-00197824.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;
XX WPI; 2003-679495/64.
XX New isolated polynucleotide related to cancer genes, useful for
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing
PT or treating cancers, e.g. breast and prostate cancers.
XX Disclosure; Fig 20; 128pp; English.
XX The invention relates to a novel isolated polynucleotide comprising a
CC differentially-regulated mammalian cancer gene. The polynucleotides of
CC the invention demonstrate cytostatic activity and are differentially
CC expressed in prostate cancer. The polynucleotide, polypeptides and
CC methods of the invention may be useful for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating cancers,
CC particularly breast and prostate cancers. Furthermore, the invention may
CC be utilised during gene therapy procedures or in the production of
CC transgenic animals. The current sequence is that of the prostate cancer-
CC related protein of the invention. The C-terminus of the current sequence
CC is illegible within the specification and as a result, residues 1143-1715
CC are absent.
XX Sequence 1142 AA;
XX Query Match 64.9%; Score 5747; DB 7; Length 1142;
XX Best Local Similarity 96.1%; Pred. No. 0;
XX Matches 1097; Conservative 25; Mismatches 19; Indels 0; Gaps 0;
QY 1 MSVLISQSVINYEENIPALKALLEKCKVDNERECQOTPLMLAAEQGNVEIVKELKN 60
Db 1 MSVLISQSVINYEENIPALKALLEKCKVDNERECQOTPLMLAAEQGNVEIVKELKN 60
QY 61 GANCNLEDLNTWTALISASKEGHIHVEELKSGASLEHDMGWTALMWACYKGRDGV 120
Db 61 GANCNLEDLNTWTALISASKEGHIHVEELKSGASLEHDMGWTALMWACYKGRDGV 120
QY 121 ELLISHGANPSVTGLQSVYPIIWAAGRGHADIHLLONGAKVNCSDKYTTPLVWAAR 180
Db 121 ELLISHGANPSVTGLQSVYPIIWAAGRGHADIHLLONGAKVNCSDKYTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVDEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGN 240
Db 181 KGHLECVKHLAMGADVDEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGN 240
QY 241 LMTASKEGHIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
Db 241 LMTASKEGHIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
QY 301 QDNKTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360
Db 301 QDNKTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360
QY 361 KVSADVKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPNYDCSHQSKIL 420
Db 361 KVSADVKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPNYDCSHQSKIL 420
QY 421 TQIFGARHLSPTETDGDMLGYDLYSSALADILSSEPTMQPPTCVGLYAQWGSKSFLLKKL 480
Db 421 TQIFGARHLSPTETDGDMLGYDLYSSALADILSSEPTMQPPTCVGLYAQWGSKSFLLKKL 480

```
Db 421 TQIFGARHLSPTEGDTGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSLKKL 480
Qy 481 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVFAFPVDNLAIAISLSFLALIYF 540
Db 481 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVFAFPVDNLAIAISLSFLALIYF 540
Qy 541 FIVYFGREGESWNWAWLSRLARHIGYBELLFKLMEFVNPPPELPEQTTKALPVRFPLF 600
Db 541 FIVYFGREGESWNWAWLSRLARHIGYBELLFKLMEFVNPPPELPEQTTKALPVRFPLF 600
Qy 601 TDYNRLLSVGGTSLAEMIAATLSDACEREFGLATRLFRVFRTEESQGGKKWKTCCPLPS 660
Db 601 TDYNRLLSVGGTSLAEMIAATLSDACEREFGLATRLFRVFRTEESQGGKKWKTCCPLPS 660
Qy 661 FVFLFVIGGIIAGITLLAIFRVDPKHLTVNAILISIASVVGAFVNLNCRTHWQVLDL 720
Db 661 FVFLFVIGGIIAGITLLAIFRVDPKHLTVNAILISIASVVGAFVNLNCRTHWQVLDL 720
Qy 721 NSQKRLHSAASKLHKLKSGFMKVLKCEVELMARWAKTIDSTFQNTQRLVWIIDGLDAC 780
Db 721 NSQKRLHSAASKLHKLKSGFMKVLKCEVELMARWAKTIDSTFQNTQRLVWIIDGLDAC 780
Qy 781 EQDKVLQMLDTPVRLFSKGFPIAFASDPHIIKAINQNLSVLRDSNININGHDYMRNIVH 840
Db 781 EQDKVLQMLDTPVRLFSKGFPIAFASDPHIIKAINQNLSVLRDSNININGHDYMRNIVH 840
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVRSQNSLGENTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVRSQNSLGENTKLGSKTALN 900
Qy 901 RRDYRRRQMTITRQMSFDLTKLVTEWDFSDISPDWRRLLNIVSVTGRLLRANQIT 960
Db 901 RRDYRRRQMTITRQMSFDLTKLVTEWDFSDISPDWRRLLNIVSVTGRLLRANQIS 960
Qy 961 FNWDLASWINLQEQMPYRTSWLILYLETEGIPDQMTLKTMYERISKNIPPTKDVEPLL 1020
Db 961 FNWDLASWINLQEQMPYRTSWLILYLETEGIPDQMTLKTMYERISKNIPPTKDVEPLL 1020
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLECTVNLDPKLEIITADYRAAEQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLECTVNLDPKLEIITADYRAAEQISIGGLAY 1080
Qy 1081 PPLPLHEGPPRPSPGVSQPVASVCSSAFNGFPFGVVSPPHSSYSGLSGPQHPFNRA 1140
Db 1081 PPLPLHEGPPRPSPGVSQPVASVCSSAFNGFPFGVVSPPHSSYSGLSGPQHPFNRA 1140
Qy 1141 A 1141
Db 1141 S 1141

RESULT 14
ID ADC06846
AC ADC06846 standard; protein; 1184 AA.
XX
AC ADC06846;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human Kidine220Pc protein XM_045362.
XX
KW cytostatic; prostate cancer; breast; gene therapy; transgenic; human;
KW Kidine220Pc; kinase D-interacting substrate of 220KDa; chromosome 2p25.1.
XX
OS Homo sapiens.
XX
PN WO2003064599-A2.
XX
PD 07-AUG-2003.
XX
PF 24-JAN-2003; 2003WO-US001943.
XX
PR 25-JAN-2002; 2002US-00054935.
```

```
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00144198.
PR 19-JUL-2002; 2002US-00197824.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;
XX
XX WPI; 2003-679495/64.
XX
XX New isolated polynucleotide related to cancer genes, useful for
XX detecting, diagnosing, staging, monitoring, prognosticating, preventing
XX or treating cancers, e.g. breast and prostate cancers.
XX
XX Disclosure; Fig 20; 128pp; English.
XX
XX The invention relates to a novel isolated polynucleotide comprising a
XX differentially-regulated mammalian cancer gene. The polynucleotides of
XX the invention demonstrate cytostatic activity and are differentially
XX expressed in prostate cancer. The polynucleotide, polypeptides and
XX methods of the invention may be useful for detecting, diagnosing,
XX staging, monitoring, prognosticating, preventing or treating cancers,
XX particularly breast and prostate cancers. Furthermore, the invention may
XX be utilised during gene therapy procedures or in the production of
XX transgenic animals. The current sequence is that of the prostate cancer-
XX related protein of the invention. The current sequence is not fully
XX legible within the specification; the complete sequence was obtained from
XX GenBank.
XX
XX Sequence 1184 AA;
```

```
Query Match 64.7%; Score 5730.5; DB 7; Length 1184;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MSVLISQSVINVEENIPALKALEKCKDVERNECGGTPLMLAAEQGNVEIVKELKN 60
Db 1 MSVLISQSVINVEENIPALKALEKCKDVERNECGGTPLMLAAEQGNVEIVKELKN 60
Qy 61 GANCNLEDLNDNTALISASKEGHIHIVBELLKSGASLEHRDMGCGTALMWACVKGRTDVV 120
Db 61 GANCNLEDLNDNTALISASKEGHVHIVEBLLKGVNLEHRDMGCGTALMWACVKGRTDVV 120
Qy 121 ELLLSHGANSPTGLQYVYPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPLVAAAR 180
Db 121 ELLLSHGANSPTGL-QYVYPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPLVAAAR 179
Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
Db 180 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 239
Qy 241 LMTASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKVIADIRG 300
Db 240 LMTASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKVIADIRG 299
Qy 301 QDNKTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 360
Db 300 QDNKTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHQKSL 420
Db 360 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHQKSL 419
Qy 421 TQIFGARHLSPTEGDTGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSLKKL 480
Db 420 TQIFGARHLSPTEGDTGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSLKKL 479
Qy 481 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVFAFPVDNLAIAISLSFLALIYF 540
Db 480 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVFAFPVDNLAIAISLSFLALIYF 539
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:18:20 ; Search time 40.5 Seconds
(without alignments)
3161.065 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSTGFGBERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	575	6.5	2753	4	US-09-949-016-7659 Sequence 7659, Ap
2	575	6.5	2753	4	US-09-949-016-7660 Sequence 7660, Ap
3	558	6.3	1883	4	US-09-949-016-9010 Sequence 9010, Ap
4	558	6.3	1883	4	US-09-949-016-9011 Sequence 9011, Ap
5	558	6.3	1883	4	US-09-949-016-9012 Sequence 9012, Ap
6	558	6.3	1883	4	US-09-949-016-9013 Sequence 9013, Ap
7	558	6.3	1883	4	US-09-949-016-9014 Sequence 9014, Ap
8	558	6.3	1883	4	US-09-949-016-9015 Sequence 9015, Ap
9	558	6.3	1883	4	US-09-949-016-9016 Sequence 9016, Ap
10	558	6.3	1883	4	US-09-949-016-9017 Sequence 9017, Ap
11	556	6.3	1856	4	US-09-949-016-6964 Sequence 6964, Ap
12	556	6.3	1880	4	US-09-949-016-5876 Sequence 5876, Ap
13	556	6.3	1881	4	US-09-949-016-5876 Sequence 5876, Ap
14	550	6.2	3924	4	US-09-538-092-1246 Sequence 1246, Ap
15	545	6.2	1719	4	US-09-949-016-6966 Sequence 6966, Ap
16	536	6.1	1839	2	US-09-172-977-4 Sequence 4, Appli
17	536	6.1	1839	4	US-09-404-108-4 Sequence 4, Appli
18	525	5.9	4377	4	US-09-949-016-6978 Sequence 6978, Ap
19	503.5	5.7	843	2	US-09-172-977-3 Sequence 3, Appli
20	503.5	5.7	843	4	US-09-404-108-3 Sequence 3, Appli
21	466.5	5.3	1745	2	US-09-031-485-33 Sequence 33, Appl
22	466.5	5.3	1745	2	US-08-847-429A-33 Sequence 33, Appl
23	466.5	5.3	1745	3	US-09-065-474-33 Sequence 33, Appl
24	466.5	5.3	1745	3	US-09-557-034-33 Sequence 33, Appl
25	454.5	5.1	3913	4	US-09-949-016-10933 Sequence 2, Appli
26	449.5	5.1	1088	3	US-09-082-059-2 Sequence 8307, Ap
27	439.5	5.0	1881	4	US-09-949-016-8307 Sequence 8307, Ap

28	434	4.9	994	4	US-10-164-595-38 Sequence 38, Appl
29	414.5	4.7	1423	3	US-08-810-712-10 Sequence 10, Appl
30	414.5	4.7	1431	4	US-09-538-092-1198 Sequence 1198, Ap
31	409.5	4.6	683	4	US-10-164-595-71 Sequence 71, Appl
32	395.5	4.5	1327	3	US-09-196-387-2 Sequence 2, Appli
33	395.5	4.5	1327	4	US-09-841-835-2 Sequence 2, Appli
34	395.5	4.5	1327	4	US-09-972-115A-8 Sequence 8, Appli
35	394	4.5	1619	4	US-09-392-812A-4 Sequence 4, Appli
36	392	4.4	1704	4	US-09-392-812A-2 Sequence 2, Appli
37	387.5	4.4	949	3	US-09-196-387-10 Sequence 10, Appl
38	387.5	4.4	949	4	US-09-841-835-10 Sequence 10, Appl
39	384.5	4.3	546	4	US-09-949-016-7169 Sequence 7169, Ap
40	380.5	4.3	1100	4	US-09-427-154-2 Sequence 2, Appli
41	380.5	4.3	1100	4	US-09-696-668-3 Sequence 3, Appli
42	380.5	4.3	1166	4	US-09-972-115A-6 Sequence 6, Appli
43	380.5	4.3	1227	4	US-09-849-602-26 Sequence 26, Appl
44	380.5	4.3	1240	4	US-09-696-668-4 Sequence 4, Appli
45	376	4.2	348	2	US-09-031-485-28 Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-7659
; Sequence 7659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7659
; LENGTH: 2753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7659

Query Match 6.5% ; Score 575 ; DB 4 ; Length 2753 ;
Best Local Similarity 21.7% ; Pred. No. 2.5e-38 ;
Matches 411 ; Conservative 292 ; Mismatches 769 ; Indels 426 ; Gaps 76 ;

Qy	21	LKALLEKCDVDERNECGQTPMLAAEQGNVEIVKLLKNGANCNLEBDLNDWTALISASK	80
Db	187	VELLLERGAFLARTKNGLSPLHMAAQGDHVECVKHLLOHKAQVDDVTLDTLTAHVAAH	246
Qy	81	EGTHIVEELLKSGASLEHDMCGWTALMWCYKGRDVTVELLSHGNCN-PSVTGLQVSV	139
Db	247	CGHYRVTKLLLDKRNPNARALNGFTPLHACKNRKIKVMELLVKYGASIOAIT--ESGL	304
Qy	140	YPIIWAAGRGHADIVHLLLNQNGKVCSDKYGTTPLVMAARKGHLECVKHLAMGADVQ	199
Db	305	TPHVAAFMGLNIVLLLNQNGASPDVTNIRGETALHMAARAGQVEVRCLLRNGALVDA	364
Qy	200	EGANSMTALIVAKGGYTSQVKEILKNPNVNLTKDGNLTALMTASKEGHEIVEVDLLDA	259
Db	365	RAREEQPLHIAASRLGKTEIVQLLQHMAHPDAATNGYTPHISAREQVDVASVLEEA	424
Qy	260	GTVYNIPDRSGDVTVLICAVRGHVEIVRALLQKVDADIRGQDNKTALYVAVEKGNATMV	319
Db	425	GAHSLATKKGFTPLHVAAKYGLSDVAKLLQRRRAADSAGKGLTPLHVAHAYDNQKVA	484
Qy	320	RDILQCNPDTEICTKDGFTPLIKATKMRNIEVVVELLLDKGAKVSAVDKKGDTPLHVAIRG	379

Db 485 L L L L K G A S P H A T A K N G Y T P L H I A A K N Q M O T A S T I L N Y G A E N I V T K G V T P L H L A S O E 544
Qy 380 R R R L A E L L R N P K O G R L L Y R P N K A G E T P Y N I D C S H O K - - - - S I L T Q I F G A R H L S T E T Y D 435
Db 545 G H T D M V T L L L - - - - D K G A N I H M S T K S G L T S I L H L A A Q B D K V N A D I L T K - H G A Q D A H T K - - 598
Qy 436 G M L G Y D L Y S S A L A D I L S E P T W O P P I C V G I Y A Q M G S G K - - S P L I K K - L E D E M K T F A G O Q T 492
Db 599 - - - - L G Y - - - - - T P L I V A C H - - Y G N V K M V N F L L K O G A N N A K T K N G Y - - 634
Qy 493 E P L F O F S - - - - W L I V F L T L L C G G L G V F A P P V D T N L A I A I S I S F L A L I Y I F F T V I Y F G R 549
Db 635 T P L H Q A A Q O O G H T I I N V L L O H G A K P N A T T A N G T A L A I A K R L C Y I S V D T L K V - - - - - 688
Qy 550 R E G E S N W A W A - S T R I A R H - I G Y L E L L F K L M - - - - - F V N P P E L P E Q T 591
Db 689 - - - - - T E E V T T T T T I T E K H K L A N P E T W T E V L D V S D E G D D T M T G D G G Y L R P E D L K E L G D 744
Qy 592 K A L P - V R F L - F T D Y N R L S S V G G E T S L A E M A T I L S D A C - - - - - E R E F G F L A 634
Db 745 D S L P S Q F L D G M N Y L R Y S L E G G R S D S R S H T L S H A S Y L R D S A V M D D S V V I P S H Q V S T L A 804
Qy 635 T R L F R V R T E S O C K K W - - - - - K T C C L P S F V I F L F I V - - - - - G C I I A G I T L 677
Db 805 K - - - - - E A R N S Y R L S W G T E N L D N V A L S S P I H S G F L V S F W D A R G G A M R G C R H N G L A I 858
Qy 678 L A I F R V D P K H L T V N A I L I S I A S V V G L A F V L N C T W Q V L D S L N S O R K R L - - - - - H 728
Db 859 I - - - - - I P P R K T A P T - - - - - R V T C R L - - - - - V K R H L A T M P P M V E G E 891
Qy 729 S A A S K L H L K S E G - - - - - F M K V L K C E V E L M A R M A K T I D S F T Q N - - - - - Q T R L 770
Db 892 G L A S R L I E V G P S G A Q F L G P V I V E I P H A A L R G K E R L V L R S E N G D S W K E H F C D Y T E D E L 951
Qy 771 V T I I D G L A C E Q D K V L Q M L D T V R L F S K G P F I A F A S D P H - - I I I K A I N Q N L N S V L R D S N 828
Db 952 N E I L N G M D - - - - - E V L D S P E D L E K R I C R I I T R D F Q Y F A V V S R I K Q - - - - - D S N 996
Qy 829 I N G H D Y M R N I V H L P V F L N S R G L N A K F L V S A T - - N G D I T C S D T T G T Q B D T D R R V S Q N S 886
Db 997 L I G P E - - - - - G G V L S S T V P V Q V A F P E G A L T K R I R V G L Q - - - - - A Q P M 1035
Qy 887 L G E W T K - L G S K T A L N R D T Y - - - R R Q M Q T I T R Q M S F D L T K L L V T E D W F S D I S P O T M R R 942
Db 1036 H S E L V K I L G N K A T F S I V T L E P R R K R F K P I T T I P V P K A S D V M L N G P G G D A P - T L R L 1094
Qy 943 L I N I V S V T G R L L R A N Q I T F N W D R L A S I N I L T - - - E Q M P Y R T S - - - - - W L I L Y L E E T E G L P 994
Db 1095 L - - - - - C S I T G G T P A - - - - - Q W E D I T G T P L T F V N E C V S F T T N V S A R F W L I - - - - - D C R Q I Q 1142
Qy 995 D Q M T L K T M Y E R I S K N I P T T K D V P L L E I D G I N F E V L S S R T P V L V A R D V K T F L P C T V N 1054
Db 1143 E S V T F A S Q V Y R E I I C V P Y - - - - - M A K F V V F A K S H D P - - - - - I E A R L R C F C M 1183
Qy 1055 L D P K L - - - - - R E I I D V A R E Q I N I G L A Y P L P L H E G P R P P S G Y S Q P A S V C S S A S F 1108
Db 1184 T D D K V D K T L E Q O E N F A E V A R S R D - - - - - V E V L E G K P I Y V D C F M L V P L T K S T R Y 1232
Qy 1109 N G P - - - - - F P G G V S P Q P H S S Y S G L S G P H P F Y N R A A V P A T G S S L L S S M T V D V V C - 1160
Db 1233 D S G T L R T I I Y E G A K I H E R P G A S S Y L Q L K H F A D L Y K D D E T E S T S V L K S H L V N E V P V L 1292
Qy 1161 - - - - - E K U R I Q E G L O Q N M P Q V C T T I K K A N I N G R V L S Q C N I D E L K E M A M N F G D H L F R S M V 1217
Db 1293 A S P D L L S E V S E M Q D L I K M - - T A I L T D V S D K A - G S I K V K E L V K A A E E E F G E - - - - - 1341
Qy 1218 L E M R S V E S Q V P D P R F L N E - N S A P V P H G S A R R S S H T E L P L T E L S S O P T Y L N F S F E E 1276
Db 1342 - P R E I V E R - - V K E D E K V N E I L S G T C T R D E S S V Q S R S R G L V E - - E E W I V S D E I E E 1396
Qy 1277 L N T L G L D E G A P R H S N S W Q S Q T R R T P S L S S L N S Q D S S I E S K L T D K V Q A E Y R D A Y R E Y I A 1336

Db 1397 A R - - - - - O K A P L E I T E Y P C V E R I D K E I K G K V E K D S T G L V N Y L T D D L A N T - C V P L P K E Q I Q 1450
Qy 1337 Q M S Q L E G G T G S T I S G R S P H S T Y Y I G O S S S G S I H S T L E O E R G - - - - - K E G E 1384
Db 1451 T V O D K A G K K C E A L A V G R S E K E - - - - - G K D I P P D E T Q S T Q K Q H K P S L G I K K P V R R K L U K E Q 1506
Qy 1385 L K O E D R K S F L - - - - - M K R G D V I D Y S S G V S T W E A S P L D P I T B E D E K S D O S G S K L L P G K K S 1440
Db 1507 K O K E G L Q A S A E K A E L K G - - - - - S S E S L G E D P G L A P - - - - - E P L P T V K A 1547
Qy 1441 S E R P S L F O D L - K L K G G L R Y O K L P S D E S T G R V Q I - - - - - T P H C S K M I R T K 1488
Db 1548 T S - P L I E E T P I G S I K D K V K A L Q R V E D E - O K G R S K L P I R V K G K E D V P K K T T T H R P H A A S P 1605
Qy 1489 R L K A K O R E C A S P - - - - - Q E H S A E P I R T F I K A K E Y L S D A L L D K D S D S G S V R S N E S S P - - - - 1540
Db 1606 S L K S E R H A P G S P S P K T E R H S T - - - - - L S S A K T E R H P P V S P S K T E K H S P V S P S 1654
Qy 1541 - N H S L H N E A A D S O L E K A N L I - - E L E D E G H S - - - - - G K R G M P H S I S L G L Q D - - - - - 1582
Db 1655 A K T E R H S P A S S S K T E K H S P V S P S T K T E R H S P V S S T K T E R H P P V S P S G K T D K R P P V P S G 1714
Qy 1583 - - - - - P I T A - - - - - R M S I C S E D K K S P S E C S L I A S P E S W P A C Q K A Y N L N R T P S T V T L N 1631
Db 1715 R T E K H P P V S P G R T E K R L P V S P S G R T D K H Q P V S T A G K T E K H L F V S P S G K T E K Q P P - - - - 1768
Qy 1632 N N T A P T N A N Q N F D I E G I R E T S Q V I L R P G P S P N P T A V Q N E N L K M A H K R S O R S S Y T R L S 1691
Db 1769 - - V S P T S K - T E R I E B T M S V R E L M K A - F O S G O D P S K H K T G L F E H K S A K Q O P E K G K V R V E 1824
Qy 1692 K D - - - - - A S E L H A A S S E S - - - - - T G F G E E R E S I 1714
Db 1825 K E K G P I L T Q R E A Q K T E N Q T I K E G Q R L P V T G T A E S K R G V 1862

RESULT 2

US-09-949-016-7660
; Sequence 7660, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7660
; LENGTH: 2753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7660

Query Match 6.5%; Score 575; DB 4; Length 2753;
Best Local Similarity 21.7%; Pred. No. 2.5e-38;
Matches 411; Conservative 292; Mismatches 769; Indels 426; Gaps 76;

Qy 21 L K A L L E K K D V B E R N E C G Q T P L M A E Q G N V E I V K E L L K N G A N C N L D N W T A L I S A S K 80
Db 187 V E L L E R G A P L A R T A N G L S P L H M A A Q G D H V E C V K H L O H K A P V D V T L D Y L T A L H V A A H 246
Qy 81 E G H I H V E I L L K S A S L E H R D M G G T A L M W A C Y K G T D V V E L L L S H G A N - P S V T G L Q Y S V 139
Db 247 C G H Y R V T K L L D K R A N P N A R A L N G F T P L H I A C K N R I K M E L L V K Y G A S I Q A I T - E S G L 304
Qy 140 Y P I I W A A G R H A D I V H L L I Q N G A K V N C S D K Y G T T P L V W A A R K G H L E C V K H L L A M G A D V D Q 199

Db 305 TPIHVAAPFGLHNLVLLLLQNGASPDVTNIRGETALHMAARGOVVVRCLLRNGALVDA 364
Qy 200 EGANSNTALIVAVGGYTSVKELIKRNPVNLTKDGNALMIAKSKGHEIIVQDLDA 259
Db 365 RAREEQTPHIASRLGKTEIVQLLQHAHPDAATNGVTPHISAREGQVDVASVLLEA 424
Qy 260 GYVNIPIRSGDVTVLIGAVRGHVEIVRALLQYADIDIRGQDNKTALYMAVEKGNATWV 319
Db 425 GAAHSLATKGTPTLHVAKYGSLOVAKLLQRRAAADSAGKNGLTPLHVAHYDNQKVA 484
Qy 320 RDILQCNPDTEICTOGTEPLIKATMRNIEVELLLKQAKVSADVKKGDTPLHVAIRG 379
Db 485 LLLLEKSGASPHATAKNGVTPHIAAKNQMQIASTLLNYGAETNIVTRQGVTPHLASQE 544
Qy 380 RSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHQK----SILTOIFGARHLSPTETD 435
Db 545 GHTDMVTL----DKGANIHMSTKGLSLHLAAQEDKVNADIUTK-HGADQDHTK-- 598
Qy 436 GDMGLDYLSALADILSEPTMQPPICVGLYAQWGSKG--SFLKK-LEDEMKTIPAGOOT 492
Db 599 ---LGY-----TELIVACH--YGNVWVNFLLAQGANVNAKTNGY-- 634
Qy 493 EPLFOFS---WLIVELTLLCGGLGVFAFPVDTWLAIAISFLALYIIFVIVYFGR 549
Db 635 TPLHQAQQGHTHIINVLLQHGAKPNATTANGNTALAIKRLGYISVVDTLKVV----- 688
Qy 550 REGESWNAWALSTRLARH-IGYELLFKLM-----FVNPPPELPEQTT 591
Db 689 ----FEVTTTTTTEKHKNVPTMTVEVLDVDSDEGGDTWTGOGGYLRLPEDLKEIGD 744
Qy 592 KALP-VREL-FTDYNRLSVGGETSLAEMIAITLSDAC-----ERFGLA 634
Db 745 DSLPSQFLDGMVNLRYSLGGRSSDRSHLSHASLYRDSAYMDDSVWIPSHQVSTLA 804
Qy 635 TRLFRVFTESQKKK-----KKTCLPSFVIFLIV-----GCIAGITL 677
Db 805 K-----BAERNRYRLSGTENLDNVALSSSPIHSGFLVSPFVMDARGAMGRCRHLRI 858
Qy 678 LAIPRVDPKHLTVNAILISIASVGLAFVLCNRTWQVLDLSLNSOKRL-----H 728
Db 859 I-----IPPRCKTAPT-----RVTCRL-----VGRHLATMPMWGE 891
Qy 729 SAASKLHLKSEG-----FMKVLKCEVELMARMAKTIDSFQNTON-----QTRL 770
Db 892 GLASRLIEVGPSSGAQFLGPVIVEIHPFAALRGKRELVLVLRSENGDSWKEHFCDYDEDEL 951
Qy 771 VVIDGLDACHQDKVLQMLDTRVULFSGGPIAIPASDPH--IIKAINQNLNSVLRDSN 828
Db 952 NEILNGMD-----EVLDSPELLEKKRICRIITRDPFQYFAVVSRKQ-----DSN 996
Qy 829 INGHYDMRNIVHLPVFLNSRGLSNARKFLVTSAT--NGDITCSDTGTGTQEDTDRVSQNS 886
Db 997 LIGPB-----GGVLSVTVPVQVAVPEGALTRIRVGLQ-----AQM 1035
Qy 887 LGEMTK--LGSKTALNRDITY--RRRQMQRTITROMSPDLTKLVTEDFWSDISQPMWR 942
Db 1036 HSELVKILGNKATSPITVLEPRRKHKPIITWIPVPKASSDVMNGFGDAP-TLRL 1094
Qy 943 LLNIVSVTGRLLRANQITFNDRLASWNLNLT---EQWPYRTS-----WLILYLBETEGLP 994
Db 1095 L---CSITGGTTPA-----QWEDITGTTPLTFVNECVSFTTNVSARFWLI-----DCRQIQ 1142
Qy 995 DQMLTKMYERISKNIPIITTKDVEPILLEIDGDIRNPEVFLSSRTPLVARDVKTPLPCTVN 1054
Db 1143 ESVTFASQVYREIICVPY-----MAXFVVFKAHSDP-----TEARLRCFCM,i183
Qy 1055 LDPKL-----REITADVRAAREQINIGGLAYPLPLHEGPPRPPSGYSQASVCSSASF 1108
Db 1184 TDDKVDKTLQEQENFAEVARSD-----VEVLEKFIYVDCFGNLVPLTKSTRY 1232
Qy 1109 NGP-----PFGGVVSPQPHSSYYSGLSGQHPFPYNAAPVATGSSLLLSMTVDVVC- 1160

Db 1233 DSGTLRTTIYEGAKIHERPGASSYLOLQKHHFADLYKDDDETESTESTSVLKSHLVNEVPVL 1292
Qy 1161 --EKLROIEGLDQNMMPQYCTTIKKNANINGRVLSCQNIDELKKEMANMFGDWHLFRSNV 1217
Db 1293 ASPDLLSEVSEMKQDLIKM--TALTUVDVSKA-GSIKVELVKAABEPEGE----- 1341
Qy 1218 LEMSVESQVVPEDPRFLNE-NSSAPVPHGSSARRSSHTLPLTELSSQTPYTLNFSPEE 1276
Db 1342 -PFEIVER--VKEDLEKVNELRSGTCRDESSVQSSRSERGLVE--EEMVIVSDEEIEE 1396
Qy 1277 LNTLGLDEGAPRHNLSWQSOTRRTPSPSSLSNSQSSIEISKLTOKVOAEVRDAYREVIA 1336
Db 1397 AR-----QKAPLEITEYPCVEVRIDKEIKGVEKXDSGLVNYLTDOLNT-CVPLPKEQLQ 1450
Qy 1337 QMSLEGGTGSTTSGRSSPHSTVYVIGOSSGGGSIHSTLEQBERG-----KEGE 1384
Db 1451 TVQDKAGKCCALAVGRSSEKE---GKDIPEDETQSTQKQHKPSLGIKKPVRLKKEQ 1506
Qy 1385 LKQEDGRKSF--MKRGDVIDYSSGVSNTNEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
Db 1507 KQKEEGLQASAEKAEKKG-----SSESLGEDPGLAP-----EPLPTVKA 1547
Qy 1441 SERPSLQFQDL-KLKGGLRYOKLPSDEDESGTGRVQI-----TPHCSNMIRTK 1488
Db 1548 TS-PLIETPTGSIKDKYKALQKRVEDB-QKGRSKLPVRKGEDVPKKTTHRPHPAASP 1605
Qy 1489 RLKAKORECASP---QEHSAEPIRTFIKAEYLSDALLDKDDSSDGSVRNNESSP--- 1540
Db 1606 SLKSERHAPGSPKTEHST-----LSSAKTERHPVSPSSKTEKHSVPSPS 1654
Qy 1541 -NHSLHNEAADDQLEKANLI--BLEDEGHS-----GKRGMPHSLSGLQD----- 1582
Db 1655 AKTERHSPASSSSKTEKHSVPSPSTKTERHSPVSTKTERHPPVSPSCKTKRPPVPSG 1714
Qy 1583 -----PIIA-----RMSICSEDKSPSECSLIASSPEESHPACQAYNLNRTSTVTLN 1631
Db 1715 RTEKHPVSPGRTKELFPVSPSGRTDHPVSTAGTKEKHLFPVSPSGKTEKQPP----- 1768
Qy 1632 NNTAPTNRANQNFQIEIGRETQVILRPGSPNPTAVQENLNKSMHAKRSQSSYTRLS 1691
Db 1769 --VSFTSK-TERIETHSVRELKMA-FQSGQDPKSKHKTGLFEHKSQKQKQKQKVRVE 1824
Qy 1692 KD-----ASELHAASSES-----TGFEERESI 1714
Db 1825 KEKGPILTQREAOXNTQNTIKRGQRLPVTGTAESKRGV 1862

RESULT 3

US-09-949-016-9010
; Sequence 9010, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9010
; LENGTH: 1883.
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9010

Query Match 6.3%; Score 558; DB 4; Length 1883;

Best Local Similarity 20.6%; Pred. No. 3.5e-37; Matches 395; Conservative 270; Mismatches 635; Indels 620; Gaps 75;

Qy	21	LKALLECKDQVDRNCGQTPLMLAAEQGNEIVIKELKNGANCNLEDNWTALISAK	80
Db	97	VRELNVYGANVNAQSQGFTPLMAAQENHLEVVKFLENGANQVATEDGFTPLAVALQ	156
Qy	81	EGH-----	83
Db	157	QGHENVVAHLINYGTRKVELPALHIAARNDTRTAAVLLQNDPNPDLVSKGTFTPLHIA	216
Qy	84	-----THIVEELKSGASLEHEDMGWGTALWACYGRDVTVELLSHGANGPVTGLQYS	138
Db	217	AHYENVLAQLLNRRGASVFTFONGITPLHIAARRGNVIMVLLLDRAQOIE-TYTKDE	275
Qy	139	VYPIIWAAGRHADIHLLQNGAKVNCSDKYGTTPPLWAARKGHLECVKHLAMGADV	198
Db	276	LTPLHCAARNGHVRISBILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVRLLIQYDAEID	335
Qy	199	-----OEGA-----NSMTALIVAKGYTQSVKEILK	225
Db	336	DITLDHLTPLHVAHCGHHRVAKVLLDKGAKPNRSLNGFTPLHIAKKNHVRVMEILLK	395
Qy	226	RNPVNLTQDGNLTALMIASKEGHEIVODLLDAGTVNIPDRSGDVTVLIGAVRGHVEI	285
Db	396	TGASIDAVTESGLTPLHVASFMGHLPVKNLLQRGASPNVSVKVTPLHMAARAGHTEV	455
Qy	286	VRALLQYADIDIRGDNKTALYWAKEGNATWVRDILQCNPTDTEICTKDGETPLIKATK	345
Db	456	AKYLLQKAKVNAKDDQPLHCAARIGHTNMVKLLLENANPNLATTAGHTPLHIAAR	515
Qy	346	MRNIEVVELLDKAKVSAVDKGTPLHVAIRGRSRLAEILLRNPCKOGLLYRPNKAG	405
Db	516	EGHVETLALLEKEASQACMTKGTPLHVAARYKVRVAELLE--RDA--HENAAG	569
Qy	406	E---TPYNIDCSHQ---KSIL-----TQIFGAR-----	427
Db	570	KNGITPLHVAHVNHLNLDIVKLLPRGGSHPSPAWNGYTPHIAAKQNEVARSLQYGG	629
Qy	428	-----HLSPTDGMGLDYLLSSALADILSEPTWQ-----PTC--	462
Db	630	SANAEVQGVTPHLAAQGEHAEMVALLSKQANGNLGNKSLTPLHLVAQEGHVPVADV	689
Qy	463	-----VGLYAOWGSGK--SPLKKLED-EMKTFAGQOOTEPLFQPSWL	501
Db	690	LIKHGMVDATRMGYTPHLVASHYGNIKLVKFLLOHQADVNAKTLGY--SPLHQAAQ	747
Qy	502	--IVFLTLLCGGLGVAFPVD-----TNLAISLSFLALIYIFFVIYFGRREGE	553
Db	748	GHTDIVTLLKNG-----ASPNEVSSDGTTPPLAIKRLGYSIVTDVLKW--	792
Qy	554	SNWAWALSTRLAHIGVLELLFKLMFVNPPELPEQTTKALPVFLFTDYNELLS-SVGE	612
Db	793	TDETSFVLVSDKHR--MSPETVDEILDVSEDEGBELIS-----PRAERDRSDVEE	843
Qy	613	TSLAEMIATLSDACERBEGFLATLFR-----VFTREE--SQGKKWKTKCCLPSFV	662
Db	844	KELLDVFKLQOVE-----SPAIPRIPCAMPETVIRSEEQASKEYDEDSLIPS--	895
Qy	663	IFLFTVGCIIAGITLLAIFRVDPKHLTVNAILISASVVGAFVNLNCRTWQVLDLSLNS	722
Db	896	-----SPATETSDNI--SPVASPVHTGFLVSF-----MVDARGGS	928
Qy	723	QKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTIDSTQNTQRLVLIIDGLDACEQ	782
Db	929	MRGSRHN-----GLRWIP-----PRCAAPTRTCRLV-----	957
Qy	783	DKVLQMLDTRVLFSGKGFIAIFASDPHIIKAINQNLNSVLRDSNNGHDYMRN-IVHL	841
Db	958	--KPKLSTPPPL-----ABEGLASRIIALGPTGAQFLSPVIVEI	996
Qy	842	PVFL-NSRGLSNARKFLVTSATNGDITCSDTTGQEDTDRRVSONSLGEMTKLSKTNAL	900

Db	997	PHFASHGRG---DRELVLVRSENGSV-----WKEHRSRYGESYLDQILN-GMDEELG	1044
Qy	901	RDYVRRQMQORTTITROMSFDTLKLVT---EDFSDISPTMRRLNIVSVTGRLLRAN	957
Db	1045	SLEELEKRVCKRIIT--TDFPLYFVIMSRUCOD-YDTIGPEG-----GSLSKVLPLV	1094
Qy	958	QITFNWDRLASWINITEQWPYRTSLILYLBETEGLPQDMTLK-----TWYER	1005
Db	1095	QATFFENAVTKRVKLAQ-----AQVPDELVTKLGNQATSPITVVEPR	1140
Qy	1006	ISK-----NIFTKDVPELLEIDGDIRNFVF-----LSSRFPVLVA	1042
Db	1141	RRKFRHPIRLPLPSPSWTDNPRDSEGGDTTSLRLLCVIGTDOAQWEDITGTTKLVA	1200
Qy	1043	RDVKTFLPCTVNLDPKLREIADVRAAREQINIGLAYPPLPLHREGPPRPSGYSQASV	1102
Db	1201	NECANF---TTNV---SARFWJSDCPTAEAVNFATLLKEL-----TAVYMAKF	1245
Qy	1103	CSSASFNQPPFGGVVSPQPHSSYYSGLSGPOHPFYNRAAVPATGSSLLSSMTVDVCEK	1162
Db	1246	VIFAKMNDPRG-----RLRCYCMTDDKDKVT	1272
Qy	1163	LRQ-----IEG-----LDQNMPOYCTTIKKAN-----INGRVL	1191
Db	1273	LEQHENFEVARSRDIEVLEGLMSLFAELSGNLVP-----VKKAAQORSFHFOSFRENLA	1327
Qy	1192	SQCNIDELKKE-----MAMNFGD-WHLFRSMVLEMRSVESQVVPEDPR-----	1233
Db	1328	MPVKVDRSREPGGSLSLFKAMKYEDTCHILCHLNIITMPPCAKSGAEDRRRTPTPLAL	1387
Qy	1234	---FLNENSSAPVPHGESARR-----SSTHELPLTELSSQTPYTNFSEELNTGLD--	1283
Db	1388	RYSILSESTPGSLSGTEQAEMKMAVISEHLGLSMAELARE-----LQFSVEDINRIVENP	1443
Qy	1284	---EGAPHSNLSWQSQTRTPSLSSLSNDSQSSIEISKLTDKVOA-----EYRD	1329
Db	1444	NSLLEQSVALLNL-WIREGQANMENLYALQSDRGEIVNVLNREGSGRQSNLKPDRRH	1502
Qy	1330	AYREYIAQMSQLEGGTGSSTISGRSPHSTYYVIGOSSGSIHSTLEQERKEGELKQSD	1389
Db	1503	TURDYLSLSPQNG--YSSLQDELLSP-----ASLGCALSSPLRADQ-----	1542
Qy	1390	GRKSFMLKRGDIVDYSSSGSVSTNEASPLDPI-----TEED---EKSDQS--GSKLLPGKS	1440
Db	1543	-----YWNEVAVLDAIPLAATEHDTMLEMSDMQVWSAGLTPLSLVT	1582
Qy	1441	SERPSIFQTDKLKGGGLRYQKLPSEDESGTGRVQITPHCSKMIRTKLKAQRECAPS	1500
Db	1583	AEDSSL-----ECSKAEDSDATGH-----EWKLEGALSSEPRGP	1616
Qy	1501	QEHSAEPIRTFIKAKELYSDA---LLDKXSDSDSGVRSNESPNSHSLHNEADDSOLEKA	1557
Db	1617	ELGSLR---LVEDDTVDSDATNGLIDLLE-QBEGORSEKLPKSGKRODDATGAGQ-DSE	1670
Qy	1558	NLIELEDEGHSGKRGKPHSLSGLODPIIARMSICSEDK-----KSPSECSLIASSPEEW	1612
Db	1671	NEVSLVSGHQGOARITHS-----PTVSQVTERSODRLQDWDADGSIYSVLODAAQGSW	1724

RESULT 4

US-09-949-016-9011
; Sequence 9011, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9011
; LENGTH: 1883
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9011

Query Match      6.3%; Score 558; DB 4; Length 1883;
Best Local Similarity 20.6%; Pred. No. 3.5e-37;
Matches 395; Conservative 270; Mismatches 635; Indels 620; Gaps 75;

Qy 21 LKALLECKDVERNECCGOTPLMAAEOGNEVEIKELKNGANCNLELDLNDWTALISASK 80
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 97 VRELNYGANVNAOSQKGFPLYMAAQENHLEVKFLLLENGANQNVATEDGFTPLVALAQ 156
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 EGH----- 83
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 157 QGHENVAHLINYGTKGKVRPLPALHIAARNDTRTAAVLLQNDPNPDLVSKTGFTPLHIA 216
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 84 -----IHIVBELKSGASLEHRDMCGWTALMWACYKGRDVTVELLSHGANSVTVGLQYS 138
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 217 AHYENLNAVQALLNRGASVNFPTQNGITPLHIAARRGNVIMVRLLLDRGAQIE-TKTDE 275
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 139 VYPIIWAAGRHADIVHLLQNGAKVNSDKYGTTPVWAARKHLECVKHLANGADV 198
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 276 LTPHCAARNGHVRISIELLDHGAPIQAKTNGLSPIHMAAQGDHLDVCLLQYDAEID 335
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 199 -----QEGA-----NSMTALIVAVKGGYVTSVKETLK 225
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 336 DITLHLTPLHVAACHGHRVAKVLLDKGAPNSPALNGFTPLHIAKKNHVRVWELLK 395
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 226 RPNVNLTDKGNALTAMTASKEGHEIVODLLDAGTYVNIPIRSGDVTVLICAVRGHVEI 285
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 396 TGASIDAVTESGLTPLHVASPMGLPIVKNLLQRGASPNVSNVETPLHMAARAGHTEV 455
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 286 VRALLOKADYDIDRGQDNKNTALYAVEKGNATVRDILQCPDTEICTKDETLIKATK 345
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 456 AKYLLQNKAKYNAKADQDTPLHCAARIGTHNMVKLLNENANPNLATAGTHTPLHIAAR 515
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 346 MRNIEVVELLDKGAKVASDKGDTPLHVAIRGRSRLAELLRNPKDGLLRPNKAG 405
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 516 EGHVETVLALLEKEASQACMTKGFPLHVAKYKVRVAELLE--RDA-----HPNAG 569
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 406 E---TPYINIDCSHQ-----KSIL-----TQIFGAR----- 427
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 570 KNGLTPLHVAVHNNLDIVKLLPRGGSPHSPAWNGYTPHIAAKQNVARSILLQYGG 629
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 428 -----HLSPTETDGMGLDYLLSSALADILSEPTMOP-----PIC-- 462
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 630 SANAESVQGVTPHLAAQEGHAEVALLSKOANGNLGKSGLTPLHVAQEGHVPADV 689
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 463 -----VGLYAOWGSGK--SFLKKLED-EMKTFAGQOETPLQFSWL 501
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 690 LIKHGVMDATRMGYTPLVASHYGNIKLVKFLHQADVNAKTKGY--SPLHQAAQ 747
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 502 --IVFLTLCCGGLGVAPVD-----TNLAISLSFLALYIPFVIYFQGRREGE 553
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 748 GHTDIVTLLKNG-----ASPNEVSSDGTTPLAIAKRLGIYSVTDVLKVW----- 792
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 554 SNNWAWALSTLARHIGYLELLPKLMFVNPELPEQTTKALPVFLPTDYNRLS-SVGGE 612
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 793 TDETSFVLVSKHR-MGFPETVDSILDVSEDEBELIS-----FKAERDRSDVDEE 843
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 613 TSLAEMIATISDACERBFGFLATLFR-----VFRTEE-SQGKKWKKTCCLPSFV 662
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 844 KELLDVFKLDQVVE-----SPAIPRCPAMPETVIRSEEQEAKSEYDEDESLIFS-- 895
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 663 IFLFVIGCIITLTAIFRVPKHLITVNAILSIASVVGFLAVLNCRTWMQVLDLSLNS 722
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
/ Sequence 9012, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: Fast-Seq for Windows Version 4.0
/ SEQ ID NO 9012
/ LENGTH: 1883
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-949-016-9012

Query Match      6.3%; Score 558; DB 4; Length 1883;
Best Local Similarity 20.8%; Pred. No. 3.5e-37; Indels 620; Gaps 75;
Matches 395; Conservative 270; Mismatches 635;

QY 21 LKALEKCKDVDERNECGQTPMLAAEQGNVEIVKELLKNGANCNLEDLNNWTALISASK 80
DB 97 VRELNVGNVNAQSKGFTPLYMAAQENHLEVVKELLENGANQNVATEDGFTPLAVALQ 156
QY 81 EGH----- 83
DB 157 QCHENVVAHLINYGTKVKRLPALHIAARNDTRTAVALLQNDPNPDVLSKTGFTPLHIA 216
QY 84 -----THIVEELKSGASLEHDMGWGTALMWACYKGRDVTVELLSHGANSVVTGLQYS 138
DB 217 AHVENLVAGLLNNGASVNFQNGITPLHIAARRGNVIMVRLLDLDRGAQIB-TKTKDE 275
QY 139 VYPIIWAAGRGHADIHLLONGAKVNCSDKYGTPPLVWAARKHLECVKHLAMGADV 198
DB 276 LTPLHCAARNGHVRISBILLDHGAPIOAKTKNGLSPIHMAAQDHLDCVRLLLQYDAEID 335
QY 199 -----QEGA-----NSMTALIVAKGTYGYSVKEILK 225
DB 336 DITLDHLTPLHVAACHGHRVAKVLLDKGAKPNSRALNGFTPLHIAACKNHRVWMLLK 395
QY 226 RNPVNLTDKGNTALMIASKEGHIIEVDLDAGTYVNIPIRSGDTPVLICAVRGGRHEI 285
DB 396 TGAVIDAVTESGLTPLHVASFMGHLPIVKLLQEGASPNVNVKVEIPLHWAARAGHTEV 455
QY 286 VRALLQYADIDIRGQNKATLYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATK 345
DB 456 AKVLLQKAKVNAKADQDTPHCAARIGHTNMVKKLLENNANPNLATTAGHTAGTPLHIAAR 515
QY 346 MRNTEVELLLDKAKVSAVDKGDTPHVAIGRSERLAELLLRNPKGRLLYRPNKAG 405
DB 516 EGHVETVALLLEKASQACMTKGFTPLHVAARYKVRVAELLE--RDA-----HPNAA 569
QY 406 E---TPYNIDCSHQ-----KSL-----TQIFGAR----- 427
DB 570 KNGITPLHVAHVHNNLDIVKLLPRGGSHPSPAWNNGYTPHIAAKQNVARSLLQYGG 629
QY 428 -----HLSTETDGMGLDYDLYSSALADILSEPTWOP-----PTC-- 462
DB 630 SANAESVQGVTPPLHAAQEGHAEMVALLSKQANGNLGNKSGLTPLHVAQEGHPVADV 689
QY 463 -----VGLYAOQWGSK--SFLKKLED-EMKTPAQQTPELPQFSWL 501
DB 690 LIKHGVMDATRMGYTPLHVASHYGNIKLVKFLQHQADVNAKTKLOY--SPLHQAAQ 747
QY 502 --IVFLTLLCGGLGVFAFPVD-----TNLATAISLSFLAIYFIPIFYGGRERGE 553
```

748 GHTDVTLLKNG-----ASPNEVSSDGTTPTLAIKRLGYISVTDVLKV----- 792

554 SNNWAWALSTRLARHIGYLELLFKLMFVNPPPELPTQTKALPVRFLFTDYNRLS-SVGGE 612

793 TDTSFVLVSDKHR-MSFPETVDEILDSEDEGEELIS-----FKAERDSRDVDEE 843

613 TSLASMIATLSDACEREGFLATRLFR-----VPTTEE-SQKKKKWKKTCCLPSFV 662

844 KELLDFVPKLDQVB-----SPAIPRIPCAMPETVIVIRSEEQEASKEYDEDSLIPS-- 895

663 IFLFIVGCIAGITLLAIFRVDPKHLTVNAILISASVVGLAFVLCRTVMQVLSLLNS 722

896 -----SPATETSDNI-SPVASPVHTGFLVSF-----MVDARGGS 928

723 QKRRLHSAASKLHKLSEGFPMKVLCEVELMARMAKTIDSTFQNTQRLVVIIDGLDACQ 782

929 MRGSRHN-----GLRVVIP-----PRCAAPTRITCRLV----- 957

783 DKVLQMLDTRVRLFSKGPPIAIFASDPHIIKAINQNLNSVLDRDSNINNGHDYMRN-IVHL 841

958 --KPKXLSTPPP-----AEEGLASRIIALGPTGAQFLSPVIVEI 996

842 PVFL-NSRGLSNARKFLVTSATNGDITCSDDTGTQEDTDRRVSONSLGEMTKLGSKTALN 900

997 PHFASHGRG--DRELVLVLRSENGSV-----WKEHRSRYGESYLDQILN-GMDEELG 1044

901 RRDVTRRRQOMQRTITFQMSFDLTLLVT---EDWFSIDISPTQMRLLNIVSVTGLLRAN 957

1045 SLEELKKEKVCRIIT--TDFFLYFVIMSLCOD-YDTIGPEG-----GSLKSLVPLV 1094

958 QITFNWDRLASWINILTEQWPYRTSWLLIYLETEGLPDQMTLK-----TMYER 1005

1095 QATFFENAVTKVKLAQ-----AQVPVDELVTKLGNQATFSPVITVEPR 1140

1006 ISK-----NIPTKDVEPILLEIDDIRNEVP-----LSRTPVLVA 1042

1141 RKFHPRIGLRPLPSPDNDPRDSGEGDTSRLRLCSVIGGTDAQWEDITGTTKLVA 1200

1043 RDVKTPCTVNLDPKREIIADVRAAREQINIGLAYPPLPLHEGPPPPPSGYSQASV 1102

1201 NECANF--TTNV--SARFWSDCPRTAEAVAFATLLYKEL-----TAVPYMAKF 1245

1103 CSSASFNGFPFGGWSQPSSHYSYGLSGPQHPFYNRAAVPATGSSLLSSMTVDVCEK 1162

1246 VIFAKMNDPREG-----RLRCYCMTDDKVDKT 1272

1163 LRQ-----LEG-----LDQNMPOYCTTIKKAN-----INGRVL 1191

1273 LEQHENFVEARSRDIEVLEGMSLFAELSGNLVP-----VKAAQOQSFHOSFRENRLA 1327

1192 SOCNIDELKE-----MAMNPGD-WHLFRSMVLEMRSVESQVVVPEDPR----- 1233

1328 MPVKVRDSSRPGGSLSLFKAMKYEDTOHILCHLNTWPPCAKSGAEDRRTPPTPLAL 1387

1234 ---FLNENSAAPVPHGESARR-----SSHTEPLTELSTQPTPLNFSFEELNTGLD-- 1283

1388 RYSILSESTPGSLSGTEQAEMKMAVISEHLGLSWAELARE---LQFSVEDINRIVENP 1443

1284 ---EGAPHSNLSWOSQTRTPSLSSINSQDSIISIKLTDKVA-----EYRD 1329

1444 NSLLEQSALLNL-WVIREGQANMENLYALQSIDRGEIVNMLEGSGQSRNLKPDRRH 1502

1330 AYREYIAQMSQLEGGTGSSTISGRSPSHSYTYIGSSSGSISHTLEQERKGEKELQED 1389

1503 TDRYLSLSPQMG--YSSLQDELLSP-----ASLGCALSSPLRADQ----- 1542

1390 GRKSFLMKRGDVIDYSSSGSVSTNEASPLDPI-----TEED---EKSQOS--GSKLLPKKS 1440

1543 -----YNNVAVLDAIPLAATAEHDTMLEMSDMQVMSAGLTFTSLVT 1582

1441 SERPSIFQTDLKLKGGGLRYQKLPSEDESGETGRVQITPHCSKWIRTKRLKAKQRECAP 1500

1583 AEDSSL-----ECSKAEDSDATGH-----EWKLEGALSEPRGP 1616

QY 1330 AYREYIAQMSOLRGSTTSIGRSSPHSTYIIGQSSGSIHSTLEQERKGEKELQED 1389
Db 1503 TDRDYSLSPSQMG--YSSIQDELLSP-----ASLGKSSPLRADO----- 1542
QY 1390 GRKSFLMKRGDVIDYSSGVSTNEASPLDPI-----TEED---EKSDQS--GSKLLPKKKS 1440
Db 1543 -----YNEVAVLDAIPLAAVTEHDTMLEWSDQVWSAGLTSLVT 1582
QY 1441 SERPSLFTQDLKKGGLRYQKLPSEDESGTRVQIITHCSKMIKTRKAKORECASP 1500
Db 1583 AEDSSL-----ECSKAEDSDATGH-----EWKLEGALSSEPRGP 1616
QY 1501 QEHSAPERTIRIKAEVLSA---LLDKDSSDSVRSNESSPHSHLNEAADDSOLEKA 1557
Db 1617 ELGSLE-----LVBDTVDVSDATNGLIDLLE--QEGORSEBKLPGSKRQDDATGAGQ--DSE 1670
QY 1558 NLIELEDEGHSGRGMPSHLSGLQDPIIARMSICSEDK-----KSPSECSLIASSPEESW 1612
Db 1671 NEVSLVSGHORGQARITHS-----PTVSQVTERSQDRLODWDADGSIYSYLQDAAGSW 1724

RESULT 7

US-09-949-016-9014
; Sequence 9014, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9014
; LENGTH: 1883
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9014

Query Match 6.3%; Score 558; DB 4; Length 1883;
Best Local Similarity 20.6%; Pred. No. 3.5e-37;
Matches 395; Conservative 270; Mismatches 635; Indels 620; Gaps 75;

QY 21 LKALLEKCKVDNERNECGQTPLMLAARQGNVEIVKELLKNGKANCNLEDLNNWTALISASK 80
Db 97 VRELIVYGANVNAQSGKFTPLYMAAQENHLEVYKFLLENGANQVATEDGFTPLAVALQ 156
QY 81 EGH----- 83
Db 157 QGHENVVAHLINYGTGKVRPLPALHIAARNDDTRTAIVLQNDPNPVLKGTGFTPLHIA 216
QY 84 -----IHIVEBKSGASLEHRDMGWTALMWACYKGRDVTVDVLLSHGANSPVTLQYS 138
Db 217 AHYENLVNAQLLNRRGASVNFPTQGTPLTHIASRRGNVIMVRLLLDRGAQIE--TKKDE 275
QY 139 VYPIIWAAGRHADIHLLONGAKVNCSDKYGTTPLVWAARKHLECVKHLLAMGADV 198
Db 276 LTPLHCAARNGHVRISILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVRLLLQYDAID 335
QY 199 -----QEGA-----NSMTALIVAVKGGVTSVKELK 225
Db 336 DITLDHUTPLHVAACHRHVRVAKVLLDKGAKPNRSLNGFTPLHIACKKHVRVWMLLK 395
QY 226 RNPVNLTDKGNLTALMIASKEGHIEIVQDILLDAGTYVNIPIRSGDFTVLIGAVRGHVEI 285

Db 396 TCASIDAVTESGLTPLHVASFMGHLPIVKNLQRGASPNVNVKVTPLHMAARAGHTEV 455
QY 286 VVALLOKYADIDIRGDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATK 345
Db 456 AKYLLQNKAKVNAKQDQDTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAR 515
QY 346 MNIEVVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRRRLAEILLRNPDKGRLLYRPNKAG 405
Db 516 EGHVEIVLALLKEASQACWKTKGFTPLHVAKYKVRVAELLLE--RDA---HFNAG 569
QY 406 E---TPYNTDCSHQ-----KSIL-----TQIFGAR----- 427
Db 570 KNGLTPLHVAHNNLDIVKLLPRGGSPHSPAWNNGYTPLHIAAKQNVARSLLQYGG 629
QY 428 -----HLSPTETDGMGLGYDLYSSALADILSEPTMQP-----PIC-- 462
Db 630 SANAESVQGTPLHLAAQEGHAEMVALLLSKQANGLNKSGLTPLHLVAQEGHVPAV 689
QY 463 -----VGLYAQMGSGK--SFLKLEKLEDE--EMKTPAGQOTTEPLFQFQSWL 501
Db 690 LIKHGMVMDATTRMGYTPPLHVASHYGNIKLVKFLQHQADVNAKTKLGY--SELHQAQ 747
QY 502 --IVFLTLLCGGLGLVFAFPVD-----TNLAISLSFLALIYIFFIVIFPGRRGE 553
Db 748 GHTDIVTLLKNG-----ASPNEVSDGTTPLAIKRLGYISVTDVLKV----- 792
QY 554 SNWAWALSTRLARHIGYLELLPKLMFVNPPPELPEOTTKALPVRPLFTDYNRLS--SVGGE 612
Db 793 TDTSFVLVSDKXHR--MSFPETVEILDVSEDEBELIS-----FKARRDSRVDDE 843
QY 613 TSAEMIATLSDACERFGLATRLFR-----VFRTEE--SQGKKWKTKCCLPSFV 662
Db 844 KELLDFVPKLDQVVE-----SPAIPRIPCAMPETVIRSEEQEQAKEYDEDSLIPS-- 895
QY 663 IFLFVGCIIAGITLAIARVDPKHLTVNAILISIASVGLAFVLCNRTWQVLDSELLNS 722
Db 896 -----SPATETSDNI--SPVASPVHTGTFLVSF-----MVDARGGS 928
QY 723 QRKRLHSAASKLHLKSEGMKVLKCEVELMARMAKTIDSTQNTQNTRLVVIDGLDACEQ 782
Db 929 MRGSRHN-----GLRVIP-----PRYCAAPRITCLV----- 957
QY 783 DKVLQMLDTRVRLFSKGPFIAPASDPHIIKAINQNLNSVLNRDSNINGHDYMRN--IVHL 841
Db 958 --KPKLSTPPPL-----AEEGLASRIIALGPTGAQFLSPVIVEI 996
QY 842 PVFL--NSRGLSNARKFLVTSANGDITCSDTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 997 PHFASHGRG---DRELVLRSNGSV-----WKEHRSRYGESYLDQILN--GMDEELG 1044
QY 901 RRDYRRRQWORTITRQMSFDLTCLKAVT---EDWFSDDISQPTWRRLLNIVSVTGRLLRAN 957
Db 1045 SLEELKKRVCRIT--TDFPLYFVMSRLCQD--YDTIGPEG-----GSLAKSLVPLV 1094
QY 958 QITFNWDRLASWINLQWPIYRTSWLILYLEETEGPLDQMTLK-----TWYER 1005
Db 1095 QATFPENAVTKRVKLAQ-----AQVPDELVTKLGNQATFSPVIVEPR 1140
QY 1006 ISK-----NIPPTKDVPELLEIDGDIRNEFV-----LSRTPVLVA 1042
Db 1141 RRKPHRPIGLIRIPLPSPWTDNPRDSGEGDTLSRLLCVSGVIGTQDAQMEDITGTTKLVA 1200
QY 1043 RDVKTFLPCTVNLDPKLEIADVRAAREQINIGGLAYPLPLHGEPPRPSPGSPASV 1102
Db 1201 NECANF--TTNV--SARFWSDCPRTAAVNFATLLYKEL-----TAVPYMAKF 1245
QY 1103 CSSASFNPGPPGVSPQPHSSYYSGLSGPHFPYNAAPVATGSSLLSSMTVDVVECK 1162
Db 1246 VIFAKVNDPREG-----RLRCYCMTDDKVDKT 1272
QY 1163 LRQ-----TEG-----LDQNMMPQYCTTIKKAN-----INGRVL 1191
Db 1273 LEQHENFVEVARSRDIEVLLEGMSLFAELSGNLVFP-----VKKAAQORSFHQSFRENRLA 1327


```

1192 QY SQCNIDELKKE-----MANVFGD-WHLFRSMVLEMRKSVSVQVVPEDPR----- 1233
1325 Db MPVKVDRSSREPGGSLSLFLRKAMKYEDTOHILCHLNITMPCCAKGSGAEARRRTPPTPLAL 1384
1234 QY ---FLNENSSAPVPHGESARR-----SSHTEPLTELSSQTPPYTLNPFSEELATGLD-- 1283
1385 Db RYSLSESTPGSLSGTQAEKMKAVISEHLGLSWAELARE-----LQFSVEDINRIRVENP 1440
1284 QY -----EGAPRHSNLSWOSQTRRTPTSLSLNSQDSSEISKLTDRKVA-----EYRD 1329
1441 Db NSLLEQSVALLNL-WVIREQONAMNLYTALQSIDRGEIVNMLEGSGRQSRNLKPDORRH 1499
1330 QY AYREYTAQMSOLEGGTGSSTISGRSSPHSTYYIGQSSSGGSIHSTLEQERKGEGLKQED 1389
1500 Db TDRDYSLSPQNMNG--YSSLQDELSP-----ASLGCALSSPLRADQ----- 1539
1390 QY GRKSFLLMKRGDVIDYSSSGVSTNEASPLDPI-----TEED---EKSQDS---GSKLLPGKKS 1440
1540 Db -----YNEVAVLDAIPLAATEHDTMLEMSDMQVWSAGLTGPSLVT 1579
1441 QY SERPSLFTQDLKUKGGGLRYQKLPSEDESGTGRVOITPHCSKMIKTRLUKAKORECASP 1500
1580 Db AEDSSL-----ECSKAEDSDATGH-----EWKLEGALSEEPRGP 1613
1501 QY OEHSASPIRTFIKAEVLSDA---LLDKKSDSDSGVRSNESSPHNSILHNEAADDSOLEKA 1557
1614 Db ELGSL-----LVEDDTVDSDATNGLIDLLE-QEESQORSEKLPGRKQDDATGAGQ-DSE 1667
1558 QY NLTELEDEGHSGRKGMPHSLSGLQDPPIIARMSCISBDK-----KSPSECSLIASSPEESW 1612
1668 Db NEVSLVSGHQGRQARITHS-----FTVSQVTERSQDRLQDWDADGSIIVSYLQDAAQGSW 1721

RESULT 14
US-09-538--092-1246
; Sequence 1246, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormat Version 0.9
; SEQ ID NO 1246
; LENGTH: 3924
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number Q01484
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number Q01485
; US-09-538--092-1246

```

```

Query Match      6.2%; Score 550; DB 4; Length 3924;
Best Local Similarity 20.4%; Pred.No. 6.7e-36;
Matches 417; Conservative 279; Mismatches 727; Indels 626; Gaps 78;

QY      14  EEENTPALKLEKCKDVERNECCGCTPLMLAAEQGNVEIVKELLKGCANCLLEDJNNT 73
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       73  KEGHVLVELLGRSSVDSATKGNLTALHIALAQAEVVKVLVEKAGININQSQNGPT 132
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY       74  ALISAKEGHHIVBELKSGAS-----LEHRDNG--- 103
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

133 PLYMAQENHIDVVKYLLENGANQSTATGFTPLAVALQOQHQAVALLENDTKGKVR 192
104 -----GWTALMWACYKGRDVEILL 123
193 LPALHIAARKDDTKSAALLQNDHNDVOSKMMVNRTTESGFTPLHIAAHYGNVNVATLL 252
124 LSHGANPSVTGLQSVVPIIWAAGRHADIVHLLQNGAKVNGSKYGTTPPLVWAARKG- 182
253 LNRGAADVFTA-RNGITPLHVASKRGNTNNVKKLLDRGGOIDAKTRDGLTPLHCAARSH 311
183 -----HLECVKHLAMGADVQEGANSMTALIV 210
312 DQVVELLLERGAPELLARTKNGLSPLHMAAQGDHVECVHLLQHKAPVDDVTLDTLALH 371
211 AVKGYTQSVKEILKRNPNVNLTKDGNALMTASKE----- 247
372 AARCGHYRVTKLLDRKANPNARALANGFTPLHACKNRKIKVMELLVKYGASIQAITESG 431
248 -----GHIEIVQDLDLAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLOKQVADID 297
432 LTPHVAAFMGHLNIVLLLLQNGASPDVWIRGETALHMAARAGQVEVVRCLLRENGALVD 491
298 IRQDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGETPLIKATKMRNIEVVELLLD 357
492 ARAREGTPLHIASRLGKTEIVQLLQHMHPDAATNGYTPPLHISAREGOVDVASVLE 551
358 KGAKVASDKGTPLHVAIRGRSRRLAELLR-----NPKD 394
552 AGAAHSLATKGTGFTPLHVAAYKSLDVAKLLQRRAAADSAGKNGLTPLHVAAHYDNQKV 611
395 GRLL-----YRPNKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPT 432
612 ALLLEKGAHPATAKNGYTPPLHIAKKNQWQASTLLNGAETNIVTKGVTPLHLASQ 671
433 ETGDML-----GYDIYSSALADILSEPTMOPPICVGLYAQMGSGKSFLLKKLEDEM-- 484
672 EGHTDMVTLLLDKGNIHMTKSLTS-----LHLAAQ-----EDKNV 710
485 -----KTFACQOTEPLPOFHWLIVFLTLILC--GGGLV-FAPVDNLAIALISLFLAL 536
711 ADILTRHGADQDAHTKLGYTPLIV-----ACHYGNVKNVFLLKQGANVNAKTKNGYTPL 765
537 -----IVFFTVIYFGRREGESMNWAWALSTRLARHIGYLELFLKLMFNPNP----- 584
766 HQAAQOQHTHIIIVLLOHGAKNATTANGNTALA--IAKGLVISVVDTLKVTEEVTTT 823
585 -----ELPQITKALPVRFLLTYNRLSVGETSLAEMIAATLSACREBFGFLA 634
824 TTTITEKHLNVPTMTVEVLDSDEGD-DTMTGDGGEYLRPELDELGDLSLPSQFLD 882
635 TRLPRVFRTEESQKKWKTCLPSFVIFLIVGCIAGITLAIAPRVDPKHLTVNAIL 694
883 GMYLRYSLGGRSDS-----LRSPSSDRSHTLSHASY 915
695 ISIASVVGAPVL-----NCKTW--QVLDSLLNSQKRLHSAASKLHL 737
916 LRDSAVDDSVVISHQVSTLAKAERNVSLSGTENLDNVALS-----SSPH-- 965
738 KSEGFMYLKVCELMARMARKTIDFTQNTQRLVVIIDGLDACQDKVLQMLDTRVLFPS 797
966 --SGFLVIFMVDAR-----GGMRGCRHGLR--IIIPPRKCTAPTTRVTCRLVKRHLAT 1016
798 KGPFIAPASDPHIIIIKAINQNLNVLSDNSNINGHYRN-IVHLPVFLNSRGLSNARKF 856
1017 MPMVVE-----GEGLASRLIEVGPSGAQFLGPVIVBIPHPAALRG--KEREL 1061
857 LVTGATNGDI-----TCSDTTGTQBDTRRVSQNSLGMTKLSGTALNRDTRRRQMR 912
1062 VLRSENGDSWKEHFCDYT-----EDELNEILN-GMDEVLDSPDELEKKRICR 1108
913 TITRQMSFDLTKLVTEDMFSDISQWNRLLNIVSVTGRLLRAN-----QITFNWRDLA 967
1109 IITRDF-----POYFAVWS--RIKQDSNLIGPEGGVLSSTVVVQVAVPEGALT 1156

QY 968 SWINL-TEOWPYRTSWLILYLBETEGLPDQMTL-----KTMYERISKNIPTTK--DVEPL 1019
Db 1157 KRIRVGLQAOQPMHSELVKKILGNKATFSPIVLEPRRRKFHKPIWTIPVPRASDV-ML 1215
QY 1020 LBIDGDIRNEVF-----LSSRTPVLVARDVTKTFLPCTVNLDPKLRIBIID 1065
Db 1216 NGFGDAPTRLRLCSITGTTTQAQWEDITGTTPLTFVNECVSF--TTNV--SAREWLLD 1270
QY 1066 VRAAREQINIGLAVPPLPLHEGPPRPSPGYQSPASVCSSASFGPPFGVSPQPHSSV 1125
Db 1271 CRQIQESVTFASQVREI-----IC-----VPMK 1296
QY 1126 YSGLSQPOHPFNRAAVPATGSSLLSSMTVDVWCEKLRQIEGLDQNMMPQYCTTIKKAN 1185
Db 1297 FVFAKSHDPIEAR-----LRCFQMTDKDKTLEQ-----QENFAEVARSDVEV 1342
QY 1186 INGR-VLSQC--NIDELKEMAMNFGDWHLFMSVLEMRSVSOVVPEDPRFLNENSSAP 1242
Db 1343 LEGKPIYVDFCNLVPKTS-----GOHHIFFFAFK-----ENRL-----PLFKVVRDITQ 1389
QY 1243 VPHGESARSSHTPELPTLSSTQPTYNFSEELNTLGLDEGAPHSNL----- 1292
Db 1390 EPCGR-----LSFMKEPXTSLGLVHQAICNLNITLPIYTKESB 1427
QY 1293 SWQSOTRRTPSLSSLSNQSISIEISKLTDKVQAEYRD-AYRYVIAQMSOLEGTGSGSTIS 1351
Db 1428 SDOEQEEDMTSEKNDTESTETSVLKSHLVNEVPVLASPDLLSEVSEMQLIKWT-- 1485
QY 1352 GRSSPHSTYIYOSSSGSGSIH-----STLEQBRKEGELKQEDGRKSLFKMGDVIDYSS 1406
Db 1486 -----AILTVDSDKAGSIKVKELVKAABEFPGEFIVER--VKEDLEKVNEL--R 1534
QY 1407 SGVSTNEASPLDP-----ITEE-----DEKSDQSGKL-----LPG 1437
Db 1535 SGCTCTDESSVQSSRSERGLVEEWIVSDELEEARQKAPLEITPYCVVEVRIDKEIG 1594
QY 1438 KKSSERPSTL--FOTD-----LKLKGG-----GLRYOK-----LPSDED 1468
Db 1595 KVEKSTGLVNYLTDDLNTCVPLPKBQLQTVQDKAGKCEALAVGRSSEKGDIPDET 1654
QY 1469 ESGTGRVQITPHCSKMIKRLKAKORECASQEHSAEPIRTFIKAKELYSDALLDKDS 1528
Db 1655 QSTQKHKPSLGIKKFVR--RLKEKQKQKEGLQASAE-----KAEK-----KKS 1699
QY 1529 SDSGVRSNESSPHSLHNEAADDQLE-----KANLIEDEGHSGKRGMPHSL 1577
Db 1700 SEESLGEDPLAPEPLTPVKATSPLEETPIGSIKDKVKALQKXVEDE-QKGRSKLPIRV 1758
QY 1578 SGLQDPIIARMSICSBDKSPSECSLIASSPEESWPACQKAYNLRNTPSTVTLNNNTAPT 1637
Db 1759 KGKED-----VPKKTTHRPHPAAS-----PSLKERHAPGSPGPKTERHSTLSS 1802
QY 1638 NРАНQFDEIEGIRETSQVILRPGSP-NPTAVQNLNLSMAHKSQRSSYTLSDKASE 1696
Db 1803 SAKTERHPVPSPSKTEK-----HSPVSPSA-----KTERHSPASSSKT-----E 1843
QY 1697 LHAASSEST 1705
Db 1844 KHSFVSPST 1852

RESULT 15

US-09-949-016-6966

; Sequence 6966, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14


```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version

```

```

; SEQ ID NO 6966
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6966

```

Query Match 6.2%; Score 545; DB 4; Length 1719;
Best Local Similarity 20.6%; Pred. No. 3.8e-36;
Matches 376; Conservative 270; Mismatches 604; Indels 574; Gaps 71;

Qy		21	LKALLECKDWDVDRNECGQTPLMLAAQGNVEIVIKELLNKGANCNLEBDLNDWTALISASK	80
Dq	:	94	VRELNVYGANVNAQSOKGFTPLWAAQENHLEVVKFLLENGANQNATEDGFPTPLAVALQ	153
Qy	:	81	EGH-----	83
Dq	:	154	OQHENVVAHLINVTGKGVKLPAHLIAARNDTRTAIALLONDPNPVLTKTGTPLHIA	213
Qy	:	84	----ITHIVEELLKSASLEHRDMGWNTALMWACYKGRTDVVELLIISHGANPSTGLQYS	138
Dq	:	214	AHYENLVAQLLNRGASVNFTPONGITPLHIASRRGNVIWVRLLDRGAIIE-TKTDE	272
Qy	:	139	VPIIWAAGRGHADIVHLLLONGAKVNCSDKYGTPTPLWAAARKHGHEVCVXHHLLAMGADV	198
Dq	:	273	LTPHCAABNGHVRISEILLDHGAPIQAOKTNGLSPTHMAAQGDHLDVCVRLLLQYDAEID	332
Qy	:	199	-----QEGA-----NSMTALIIVAKGYTOSVKEILK	225
Dq	:	333	DITDLHLTPLHVAACHGHRVAKVLDDKGAKPNSRALNGFTPLHIACKNHVRVMELLKK	392
Qy	:	226	RNPVNLTDKDGNTALMIASKEGHIPIVDLLDAGTYVINIPDRSGDTVLIGAVRGGHVEI	285
Dq	:	393	TGASIDAVTESGLTPLHWASFMGHLPVKNLLQRGASPVSNNKVETPLHMAARAAGHTEV	452
Qy	:	286	VRALLOKYADIDIRGOONTALYAWEKGNATWVRDILOCPNPDTEICTKDGEPTLKATK	345
Dq	:	453	AKYLLONKAKVNAKADDOTPLHCAARIHTNNWKULLENANPNLATTAGHTPLHIAAR	512
Qy	:	346	MNIEVVELLDDKGAKVSAVDKKGDTPLHVAIRGRGRRLAEILLRNPKDGRLLYRPNKAG	405
Dq	:	513	EGHVETVTLALLEKEASAOCMTKGGFTPLHVAAYKGVRAELLLE-RDA---HPNAAG	566
Qy	:	406	E--TPYNIDCSHQ--KSII-----TOIFGAR-----	427
Dq	:	567	KNGLTPLHVAVHNHNLIDVKLLPRGSPHSPPAWNVTPLHIAAKONQVARSLLQYGG	626
Qy	:	428	-----HLSPTEDGMGLGVOLYSSALADILSPTWOP-----PIC--	462
Dq	:	627	SANAESVOGVTPLHIAAQEGAEMVALLISKQANGNLGNKSGITPLHVAQEGHVPADV	686
Qy	:	463	-----VCLYAQWSGK--SFLLKLED-EMKTFAGOOTPELFOFSWL	501
Dq	:	687	LKHGWVMDATTMRGYTPLHVAASHYGNIKLVLKQLHQADVNAKTKLGY-SPLHQAAQQ	744
Qy	:	502	--IVFLTLLCGLGLVFAPVD-----TNLAIALSLFLALIYFFVIYGGREGE	553
Dq	:	745	GHTDIVTLLKNG-----ASPNEVSSDGTPTPLAIAKRLGISVTDVLKV-	789
Qy	:	554	SNWNWALSTRLARHIGYLELFLMFVNPPPELTQKALPVRFIFTDYNRLS-SVGGE	612
Dq	:	790	TDETSFIUVDKXR-MSFPETVDEILDVSEDEEELIS-----FKARRRSRDVDEE	840
Qy	:	613	TSLAEMIATLSDACEREFGLATRLFR-----VFRTEE-SQGKKKWKTCCPLPSFV	662
Dq	:	841	KELLDVPKLDOWE-----SPAIPICAMPETVVISREEOBASKEYDEDSLIPS--	892

Search completed: February 5, 2005, 19:35:45
Job time : 71.5 secs

Qy	663	IFLFIWCIIGIATITLLAIFRVDPKHITVNAILLISIASVVLGAFVFLNCRNTWQVLDLSLNS	722
Db	893	-----SPATETSDNI--SPVASPVHTGLFVSF-----MVDARGGS	925
Qy	723	QRKRLHSAASKLHLKXSEGFMKVLKCEVELMARMATIDSFQCNQOFLAVIIDGLDACEQ	782
Db	926	MGRSRHN-----GLRVVIP-----PRTCAAPRITCRLV-----	954
Qy	783	DKVLQMLDITVRVLFSKGPFIAPASDPHHIIKAINQNLNSVLDRNSINGHIDYMRN--IVHL	841
Db	955	--KPQKLSPPPL-----AEEGLASRIIALTGPTGAQFLSPVIVEI	993
Qy	842	PVEL--NSRGLSNARKFLVTSATNGDITCSDTGTQBDTDRRVSONSLGEMTKLGSKTALN	900
Db	994	PHFASHRG---DRELVILRSENGSV-----WKEHRSRYGESYLDQILN--GMDDELG	1041
Qy	901	RRDYRRRQWORTITROMSFDITKLAVT--EDWFSDISPQTMRRLLNIVSVTVGRLLRAN	957
Db	1042	SLEELEKRVCRIT--TDFPLYFVIMSRLCD--YDIIGPEG-----GSLKSKLVLPLV	1091
Qy	958	QITFNWDRLASINLTQOMPYRTSMILLYLEETEGLPQWTKL-----TWYER	1005
Db	1092	QATFPENAVTKRVKALQ-----AQVPDELVTKLGNQATFSPIVTVBPR	1137
Qy	1006	ISK-----NIPTKDVEPLEIDGDIRPEVF-----LSRTPVLVA	1042
Db	1138	RRKFHRPIGLRIPLPSWTDNPRDGSGETTSLRLCSVIGGTDQAWEDDITGTTKLVA	1197
Qy	1043	RDVKTEPLCTVNLDPKRLRIIADVRAAREQINIGLAYPPLPHHEGPPRPSPGYSQPASV	1102
Db	1198	NECANF---TTNV--SARFWLSDCPTAAVNAFATLLYKEL-----TAVPYMAKF	1242
Qy	1103	CSASFNPGFPFGVWSPQPHSSYSGLSGQHPFYNNRAAVPATGSSLLSSMTVDVVCEK	1162
Db	1243	VIPAKNDPREG-----RLRCYCMTDDKVDKDT	1269
Qy	1163	LQO-----IEG-----LDONMPOYCTTIKKAN-----INGRVL	1191
Db	1270	LEOHENFVEVARSRDIEVLEGMSLFAELSGNLVP-----VKAAQORSFHQSFRENRLA	1324
Qy	1192	SOQNIDELKKE-----MAMNFGD--WHLFRSMVLEMRSVESQVVPEDPR-----	1233
Db	1325	MPVKVRDSSREPFGSLSLRKAMKYEDTQHILCHLNIITMPPCAKGSGAEDRRRTPTPLAL	1384
Qy	1234	----FLNENSAPVPHGESARR-----SSHTEIPLTELSSQTPYTLNFSPEELNTLGLD--	1283
Db	1385	RYSILSESTFGSLGTEQAEQMAVISEHGLSWAELARE---LQFSVEDINRIRVENP	1440
Qy	1284	---ECAPRHSNLWSQSORTRTPSLSSLSQDSSIBISKLTKVQA-----EYRD	1329
Db	1441	NSLLEQSALLNL--WIRSEQANWENLTALQSDRGIEVNWLEGGSRQSNLKPDRRH	1499
Qy	1330	AYREYIAQMSQLEGG-----TGSSTISGRSPHSTHYIGOSSGGSIHSTLEQERGKE	1382
Db	1500	TDRDYSLSQMGHGORGQARITHSPTVS--QVTERSQDRLQWDADGGSIVSLQD--AAQ	1556
Qy	1383	GELKQO--DGRKSF-----MKRGDVIDYSSGCVSTNEASPLDPITEDEKSDQSGS	1432
Db	1557	GSWQEEVWTQPHSFQGTSTWTEGLEPGGSOEYKVLVSVSEHTWTQPEAESQADRRR	1616
Qy	1433	KLLPGKSSERPSTFOTDLKGGGLRYOKLPSDE-----DESGTRGVQITPHCSKMI	1485
Db	1617	QOQOEBOVEAKNTFTQVVQ---GNEFQNI FGEQVTEBQFTDEQGN-----I	1660
Qy	1486	RYTKRLKAK---ORECAS---POEH	1503
Db	1661	VTKKIIRKVVROIDLSSADAQOEH	1684

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:28:23 ; Search time 141 Seconds
(without alignments)
3961.558 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSTGGERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8853	100.0	1715	14	US-10-021-571-2
2	8853	100.0	1715	14	US-10-117-229-6
3	8639.5	97.6	1762	14	US-10-117-229-7
4	8266	93.4	1715	14	US-10-021-571-4
5	8266	93.4	1715	14	US-10-117-229-2
6	8266	93.4	1715	16	US-10-473-574-26
7	8212	92.8	1771	14	US-10-117-229-4
8	8212	92.8	1771	14	US-10-117-229-11
9	5730.5	64.7	1184	14	US-10-117-229-3
10	5730.5	64.7	1184	14	US-10-117-229-9
11	2961.5	33.5	705	14	US-10-106-698-6378
12	2782.5	31.4	551	9	US-09-835-788A-17
13	2782.5	31.4	551	14	US-10-175-042-17

14	2376.5	26.8	1498	14	US-10-021-571-8	Sequence 8, Appli
15	2260	25.5	513	14	US-10-149-819-9	Sequence 9, Appli
16	1848.5	20.9	1398	14	US-10-021-571-6	Sequence 6, Appli
17	549	6.2	109	10	US-09-986-480-371	Sequence 371, App
18	538.5	6.1	1724	9	US-09-964-899-43	Sequence 43, Appli
19	527.5	6.0	657	15	US-10-104-047-2529	Sequence 2529, Ap
20	506.5	5.7	1330	15	US-10-108-260A-3237	Sequence 3237, Ap
21	483.5	5.5	1762	14	US-10-205-194-117	Sequence 117, App
22	479.5	5.4	1188	16	US-10-408-765A-1311	Sequence 1311, Ap
23	456	5.2	747	15	US-10-094-749-1924	Sequence 1924, Ap
24	454.5	5.1	3913	15	US-10-334-143-45	Sequence 45, Appli
25	438	4.9	740	9	US-09-835-788A-12	Sequence 12, Appli
26	438	4.9	740	14	US-10-175-042-12	Sequence 31, Appli
27	438	4.9	1569	15	US-10-275-595A-31	Sequence 4122, Ap
28	434.5	4.9	919	15	US-10-108-260A-4122	Sequence 343, App
29	431.5	4.9	1053	15	US-10-291-172-343	Sequence 343, App
30	431.5	4.9	1053	15	US-10-221-278-343	Sequence 38, Appli
31	414.5	4.7	1431	15	US-10-352-684A-38	Sequence 5, Appli
32	414.5	4.7	1431	15	US-10-045-400C-5	Sequence 1235, Ap
33	414.5	4.7	1431	15	US-10-295-027-1235	Sequence 70, Appli
34	401.5	4.5	426	9	US-09-908-711-70	Sequence 6, Appli
35	400	4.5	367	15	US-10-250-613-6	Sequence 2, Appli
36	395.5	4.5	1327	9	US-09-841-835-2	Sequence 8, Appli
37	395.5	4.5	1327	10	US-09-972-115A-8	Sequence 4, Appli
38	395.5	4.5	1327	14	US-10-199-937-4	Sequence 3, Appli
39	394.5	4.5	765	14	US-10-128-174-3	Sequence 34, Appli
40	394.5	4.5	765	14	US-10-128-174-34	Sequence 35, Appli
41	394.5	4.5	765	14	US-10-128-174-35	Sequence 36, Appli
42	394.5	4.5	765	14	US-10-128-174-36	Sequence 37, Appli
43	394.5	4.5	765	14	US-10-128-174-37	Sequence 38, Appli
44	394.5	4.5	765	14	US-10-128-174-38	Sequence 39, Appli
45	394.5	4.5	765	14	US-10-128-174-39	

ALIGNMENTS

RESULT 1

US-10-021-571-2
; Sequence 2, Application US/10021571
; Publication No. US20030168056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Haeyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI
; FILE REFERENCE: CHAOLIA
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-021-571-2

Query Match	100.0%;	Score 8853;	DB 14;	Length 1715;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1715;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAAEGQNVIVKELKN	60	
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAAEGQNVIVKELKN	60	
Qy	61	GANCNLEDLNDWTALISASKEGHIHIVBEELKKSASLEHRDMGGWTALMWACYKGRDGV	120	
Db	61	GANCNLEDLNDWTALISASKEGHIHIVBEELKKSASLEHRDMGGWTALMWACYKGRDGV	120	
Qy	121	ELLLSHGANSFVTLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYTTPLVWAAR	180	
Db	121	ELLLSHGANSFVTLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYTTPLVWAAR	180	

Db 121 ELLSHGANPSVTGLQYSVYPIIWAAGRGHADIHVHLLQNGAKVNCSDKYGTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNATA 240
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNATA 240
QY 241 LMTASKEGHIIEIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKQVADIDIRG 300
Db 241 LMTASKEGHIIEIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKQVADIDIRG 300
QY 301 QDNKALYAVEKGNATMRDILQCNPDTEICTKDGETPLIKATKRNIEVBVLLDKGA 360
Db 301 QDNKALYAVEKGNATMRDILQCNPDTEICTKDGETPLIKATKRNIEVBVLLDKGA 360
QY 361 KVSADVKKGTPLHVAIRGRSRLAELLIRNPKDGRLLYRPNKAGETPPYINIDCSHQKSL 420
Db 361 KVSADVKKGTPLHVAIRGRSRLAELLIRNPKDGRLLYRPNKAGETPPYINIDCSHQKSL 420
QY 421 TQIFGARHLSPTEITDGMGLVLYSSALADILSEPTMQPPICVGLYAQWGSFKSLKKL 480
Db 421 TQIFGARHLSPTEITDGMGLVLYSSALADILSEPTMQPPICVGLYAQWGSFKSLKKL 480
QY 481 EDEKMTFAGQOTPLFOFQSWLIVFLTLCCGLGLVFAFPVDNTNLAIAISLSFLALIYIF 540
Db 481 EDEKMTFAGQOTPLFOFQSWLIVFLTLCCGLGLVFAFPVDNTNLAIAISLSFLALIYIF 540
QY 541 FIVYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPELPPEOTTKALPVRFIF 600
Db 541 FIVYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPELPPEOTTKALPVRFIF 600
QY 601 TDYNRLSSVGGESLAEMIATLSDACEREFGLATRLFRVFRTEESQKKWKTKCLPS 660
Db 601 TDYNRLSSVGGESLAEMIATLSDACEREFGLATRLFRVFRTEESQKKWKTKCLPS 660
QY 661 FVIFLFTVGCIIAGITLLAIFRVPDPKHLTVNAILISIASVVGGLAFVLCRTWQVLDLSL 720
Db 661 FVIFLFTVGCIIAGITLLAIFRVPDPKHLTVNAILISIASVVGGLAFVLCRTWQVLDLSL 720
QY 721 NSQKRLHSAASKLHLKSGFVKVCEVELMARMAKTIDSTQNTQRLVLIIDGLDAC 780
Db 721 NSQKRLHSAASKLHLKSGFVKVCEVELMARMAKTIDSTQNTQRLVLIIDGLDAC 780
QY 781 EQDKVLQMLDTRVRLFSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHYDMRNIVH 840
Db 781 EQDKVLQMLDTRVRLFSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHYDMRNIVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSIGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSIGEMTKLGSKTALN 900
QY 901 RRDYREROMORTITROMSFDLTLLVTEWDFSDISPTMRRLINIVSVTGRLLRANQIT 960
Db 901 RRDYREROMORTITROMSFDLTLLVTEWDFSDISPTMRRLINIVSVTGRLLRANQIT 960
QY 961 FNWDLASWINLTBQWYRPSWLLIYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWINLTBQWYRPSWLLIYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIINFVFLSSRTPVLVARDVTKPLCTVNLDPKLEIITADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIINFVFLSSRTPVLVARDVTKPLCTVNLDPKLEIITADVRAAREQINIGGLAY 1080
QY 1081 PPLPLHGPPRPPSGYQOPASVCSASFNGFPFGVVSPOPHSSYSGLSGPOHPFNRA 1140
Db 1081 PPLPLHGPPRPPSGYQOPASVCSASFNGFPFGVVSPOPHSSYSGLSGPOHPFNRA 1140
QY 1141 AVPATGSSLLISSMTVDVCEKLRQIEGLQDNMMPOYCTTIKKANINGRVLSCQNIIDELK 1200
Db 1141 AVPATGSSLLISSMTVDVCEKLRQIEGLQDNMMPOYCTTIKKANINGRVLSCQNIIDELK 1200
QY 1201 KEWAMNFGDWHLFRSMVLEMRSVESQVVPDPRFLNENSGAPVPHGESARRSHTLPLT 1260
Db 1201 KEWAMNFGDWHLFRSMVLEMRSVESQVVPDPRFLNENSGAPVPHGESARRSHTLPLT 1260

QY 1261 ELSSQTPYTLNPSFEELNTLGLDEGAPRHSNLSWQSQTRRTTSLSSINQDSSIEISKLT 1320
Db 1261 ELSSQTPYTLNPSFEELNTLGLDEGAPRHSNLSWQSQTRRTTSLSSINQDSSIEISKLT 1320
QY 1321 DKVQAEYRDAYREYIAQMSQLEGGTGSGSTISGRSSPHSTYYIQSSSGSGSIHSTLQERG 1380
Db 1321 DKVQAEYRDAYREYIAQMSQLEGGTGSGSTISGRSSPHSTYYIQSSSGSGSIHSTLQERG 1380
QY 1381 KEGELKQEGGRKSFMLKRGDVIDYSSSGVSTWEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
Db 1381 KEGELKQEGGRKSFMLKRGDVIDYSSSGVSTWEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
QY 1441 SERPSLFQTDLKLKGGGLRYQKLPSEDESGTGRVOITPHCSQWITTKLKAQKQECASP 1500
Db 1441 SERPSLFQTDLKLKGGGLRYQKLPSEDESGTGRVOITPHCSQWITTKLKAQKQECASP 1500
QY 1501 QEHSAPIRTFIKAKAYLSDALDKKSDSDSGVRSNESSPNHSLHNEAADDSDQLEKANLI 1560
Db 1501 QEHSAPIRTFIKAKAYLSDALDKKSDSDSGVRSNESSPNHSLHNEAADDSDQLEKANLI 1560
QY 1561 ELEDEHSGKRGMPHSLGLQDPIIARMSICSEDKKSPSECSLIASSPESWPACOKAYN 1620
Db 1561 ELEDEHSGKRGMPHSLGLQDPIIARMSICSEDKKSPSECSLIASSPESWPACOKAYN 1620
QY 1621 LNRTPSTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPSPNPNTAVQNLNLSMAHK 1680
Db 1621 LNRTPSTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPSPNPNTAVQNLNLSMAHK 1680
QY 1681 RSQRSSYTRLSDASELHAASSESTGFGEERSIL 1715
Db 1681 RSQRSSYTRLSDASELHAASSESTGFGEERSIL 1715

RESULT 2

US-10-117-229-6
; Sequence 6, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-6

Query Match 100.0%; Score 8853; DB 14; Length 1715;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVLISQSVINYYEENIPALKALLECKDKVDNERNECGQTPLMLAAEQGNVEIVKELLKN 60
Db 1 MSVLISQSVINYYEENIPALKALLECKDKVDNERNECGQTPLMLAAEQGNVEIVKELLKN 60
QY 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120
QY 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIHVHLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIHVHLLQNGAKVNCSDKYGTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNATA 240
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNATA 240
QY 241 LMTASKEGHIIEIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKQVADIDIRG 300

```
Db 241 LMIASKEGHEIIVQDLLDAGTYVNIIPDRSGDVTLVIGAVRGHVEIVRALLQYADIDIRG 300
Qy 301 QDNKTALYAWVEKGNATVRDILOCNPDTEICTKDGETPLIKATKORNIIEVVELLDKGA 360
Db 301 QDNKTALYAWVEKGNATVRDILOCNPDTEICTKDGETPLIKATKORNIIEVVELLDKGA 360
Qy 361 KVSADVKKDGTPLHVAIRGRSRLAEALLRNPCKGRLLYRPNKAGETPNIDCSHQKSL 420
Db 361 KVSADVKKDGTPLHVAIRGRSRLAEALLRNPCKGRLLYRPNKAGETPNIDCSHQKSL 420
Qy 421 TQIFGARHLSPTETDGMGLDYDLSSALADILSEPTMQPPICVGLYAWGSGKSFLLKKL 480
Db 421 TQIFGARHLSPTETDGMGLDYDLSSALADILSEPTMQPPICVGLYAWGSGKSFLLKKL 480
Qy 481 EDEMKTFAGQQTPELPFQFSWLIIVFLTLCCGLGLVAFPPVDTNLAISLSFALIIYIF 540
Db 481 EDEMKTFAGQQTPELPFQFSWLIIVFLTLCCGLGLVAFPPVDTNLAISLSFALIIYIF 540
Qy 541 FIVYFGRREGESNNAWALSTRARHIGYLELLFKLMFVNPPPELPQTTKALPVRFLF 600
Db 541 FIVYFGRREGESNNAWALSTRARHIGYLELLFKLMFVNPPPELPQTTKALPVRFLF 600
Qy 601 TDYNRLSSVGGETSLAEMIATLSACEREFGLATRLFRVFRTEBSQKKKKWKTCCCLPS 660
Db 601 TDYNRLSSVGGETSLAEMIATLSACEREFGLATRLFRVFRTEBSQKKKKWKTCCCLPS 660
Qy 661 FVIFLFIAGCIIAGITTLAIIFRVPKHLTVNAIISIASVVGGLAVFLNCRMTWQVLDL 720
Db 661 FVIFLFIAGCIIAGITTLAIIFRVPKHLTVNAIISIASVVGGLAVFLNCRMTWQVLDL 720
Qy 721 NSQRKRLHSAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFQNTQRLVVIIDGLDAC 780
Db 721 NSQRKRLHSAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFQNTQRLVVIIDGLDAC 780
Qy 781 EQDKVQLMDTVRVLFPSGPIAFIPASDPHIIIIKAINONLSVLRDSNINGHDYMRNIVH 840
Db 781 EQDKVQLMDTVRVLFPSGPIAFIPASDPHIIIIKAINONLSVLRDSNINGHDYMRNIVH 840
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTGTQEDTDREVSONSGEMTKLASKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTGTQEDTDREVSONSGEMTKLASKTALN 900
Qy 901 RRDYTRRQMQRTITRQMSFDLTKLLVTEDFWFSDISPQTMERLLNIVSVTGRLLRANQIT 960
Db 901 RRDYTRRQMQRTITRQMSFDLTKLLVTEDFWFSDISPQTMERLLNIVSVTGRLLRANQIT 960
Qy 961 FNVDRLASWINLTQWVPYRTSWLILYLBETEGLPDQWTLKTMRYERISKNIIPPTTKDVBPLL 1020
Db 961 FNVDRLASWINLTQWVPYRTSWLILYLBETEGLPDQWTLKTMRYERISKNIIPPTTKDVBPLL 1020
Qy 1021 BIDGDIRNFVFLSSRTPVLVARVDKTPLPCTVNLDPKREIIADVRAAREQINIGGLAY 1080
Db 1021 BIDGDIRNFVFLSSRTPVLVARVDKTPLPCTVNLDPKREIIADVRAAREQINIGGLAY 1080
Qy 1081 PPLPLHEGPPRPPSGYSPASVCSASFNGPFGVSPQPHSSYVSGLSGPQHPFYNRA 1140
Db 1081 PPLPLHEGPPRPPSGYSPASVCSASFNGPFGVSPQPHSSYVSGLSGPQHPFYNRA 1140
Qy 1141 AVPATGSSLLSSMTVDVVCSEKLQIEGLDQNMMPQYCTTIKKANINGRVLSCQNBDELK 1200
Db 1141 AVPATGSSLLSSMTVDVVCSEKLQIEGLDQNMMPQYCTTIKKANINGRVLSCQNBDELK 1200
Qy 1201 KEMANNFGDHLFRSMVLEMSVSVQVPPDRFLNENSSAPVPHGESARSSHTPLPT 1260
Db 1201 KEMANNFGDHLFRSMVLEMSVSVQVPPDRFLNENSSAPVPHGESARSSHTPLPT 1260
Qy 1261 ELSSQTPYTLNFSPEELNTLGLDGAPRHSNLSQSQTRTPSLNSQDSSIEISKLT 1320
Db 1261 ELSSQTPYTLNFSPEELNTLGLDGAPRHSNLSQSQTRTPSLNSQDSSIEISKLT 1320
Qy 1321 DKVQAEYRDVREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGOSSSGSIHSTLEQERG 1380
```

```
Db 1321 DKVQAEYRDVREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGOSSSGSIHSTLEQERG 1380
Qy 1381 KEGELKQEDGRKSFIMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGRKSFIMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
Qy 1441 SERPSLFQTDLKLKGGGLRYQKLPDDESGTGRVQITPHPCSMMIRTKRLKAKQRECA 1500
Db 1441 SERPSLFQTDLKLKGGGLRYQKLPDDESGTGRVQITPHPCSMMIRTKRLKAKQRECA 1500
Qy 1501 QEHSAPRIRPIKAKYELSDALLDKDSSDGVSRNESSPHSLHNEAADDSQLEKANLI 1560
Db 1501 QEHSAPRIRPIKAKYELSDALLDKDSSDGVSRNESSPHSLHNEAADDSQLEKANLI 1560
Qy 1561 ELEDEGHSGKGMPSHLSGLQDPIIARMISCEKSPSECSLIASSPEESWPACQKAYN 1620
Db 1561 ELEDEGHSGKGMPSHLSGLQDPIIARMISCEKSPSECSLIASSPEESWPACQKAYN 1620
Qy 1621 LNRTPTSTVTLNNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHK 1680
Db 1621 LNRTPTSTVTLNNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHK 1680
Qy 1681 RSQSSSTRLSKDASELHAASSESTGFEERESIL 1715
Db 1681 RSQSSSTRLSKDASELHAASSESTGFEERESIL 1715
```

RESULT 3

```
US-10-117-229-7
; Sequence 7, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1762
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-7
```

```
Query Match 97.6%; Score 8639.5; DB 14; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;
```

```
Qy 1 MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAAEQGNVEIVKELIKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAAEQGNVEIVKELIKN 60
Qy 61 GANCNLEDLNTWLTALISASKEGHTHIVEELLKSGASLEHRDMGWTALMWACYKGRTDVV 120
Db 61 GANCNLEDLNTWLTALISASKEGHTHIVEELLKSGASLEHRDMGWTALMWACYKGRTDVV 120
Qy 121 ELLILSHGANPSVTGLQYSVYPIIWAAGRGHADI VHALLONGAKVNCSDKYCTTFLVWAAR 180
Db 121 ELLILSHGANPSVTGLQYSVYPIIWAAGRGHADI VHALLONGAKVNCSDKYCTTFLVWAAR 180
Qy 181 KGHLECVKHLAMGADVQEGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240
Db 181 KGHLECVKHLAMGADVQEGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240
Qy 241 LMTASKEGHEIIEVDLLDAGTYVNIIPDRSGDVTLVIGAVRGHVEIVRALLQYADIDIRG 300
Db 241 LMTASKEGHEIIEVDLLDAGTYVNIIPDRSGDVTLVIGAVRGHVEIVRALLQYADIDIRG 300
Qy 301 QDNKTALYAWVEKGNATVRDILOCNPDTEICTKDGETPLIKATKORNIIEVVELLDKGA 360
Db 301 QDNKTALYAWVEKGNATVRDILOCNPDTEICTKDGETPLIKATKORNIIEVVELLDKGA 360
```

QY 361 KVSADVKKGDTPLHVAIRGRSRELABELLRLNPKDGLLYRPNKAGETPNIDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHVAIRGRSRELABELLRLNPKDGLLYRPNKAGETPNIDCSHOKSIL 419
QY 421 TQIFGARHLSPTTGDGMLGYDYSALADIISEPTWQPPICVGLYAQWGSKGKFLKKL 480
Db 420 TQIFGARHLSPTTGDGMLGYDYSALADIISEPTWQPPICVGLYAQWGSKGKFLKKL 479
QY 481 EDEWKTFAQOQTEPLFOFQSWLIVFLTLILCGGLGLVFAFPVDTNLAIALISLFLALIIYIF 540
Db 480 EDEWKTFAQOQTEPLFOFQSWLIVFLTLILCGGLGLVFAFPVDTNLAIALISLFLALIIYIF 539
QY 541 FIVIFGGRREGESWNWALSTRLARHIGYLLFKLFPVNPPELPEQTTKALPVRFPLF 600
Db 540 FIVIFGGRREGESWNWALSTRLARHIGYLLFKLFPVNPPELPEQTTKALPVRFPLF 599
QY 601 TDYNRLSSVGGTSLAEMIAATLSDACEREFGLATRLFRVFRTEESQKKWKKTCCPLPS 660
Db 600 TDYNRLSSVGGTSLAEMIAATLSDACEREFGLATRLFRVFRTEESQKKWKKTCCPLPS 659
QY 661 FVIFLFTVGCIIAGITLILAFRVPDKHLTVNAILISIASVVGGLAFVNLNCRWTWQVLDL 720
Db 660 FVIFLFTVGCIIAGITLILAFRVPDKHLTVNAILISIASVVGGLAFVNLNCRWTWQVLDL 719
QY 721 NSQKRHLHSAASKLHKLKSGFMKVLKCEVELMARMAKTIDSFTQOTRLVAVIIDGLDAC 780
Db 720 NSQKRHLHSAASKLHKLKSGFMKVLKCEVELMARMAKTIDSFTQOTRLVAVIIDGLDAC 779
QY 781 EODKVLQMDIVRVLFKSGPPIAFASDPHIIKAINQNLNSVLNDSNINGHYDMRNIVH 840
Db 780 EODKVLQMDIVRVLFKSGPPIAFASDPHIIKAINQNLNSVLNDSNINGHYDMRNIVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSOTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSOTGTQEDTDRRVSONSLGEMTKLGSKTALN 899
QY 901 RBDTYRRQWORTITROMSPDLTKLLVTEDFWSDISQPTWRLNINVSVTGRILLRANQIT 960
Db 900 RBDTYRRQWORTITROMSPDLTKLLVTEDFWSDISQPTWRLNINVSVTGRILLRANQIT 959
QY 961 FNWDLASWINLTQOWPYRTSWLILYLEETEGPLDQMTLKTMYERISKNIPPTKQVBP 1020
Db 960 FNWDLASWINLTQOWPYRTSWLILYLEETEGPLDQMTLKTMYERISKNIPPTKQVBP 1019
QY 1021 BIDGIRNFVPLSRTPVLVARDVKTFPLCTVNLDPKREIIADVPAAREQINIGGLAY 1080
Db 1020 BIDGIRNFVPLSRTPVLVARDVKTFPLCTVNLDPKREIIADVPAAREQINIGGLAY 1079
QY 1081 PPLPLHEGPPRPPSGYSOPASVCSSASPNPGVSPQPHSSYVYSGLSGQPHFFYNR - 1139
Db 1080 PPLPLHEGPPRPPSGYSOPASVCSSASPNPGVSPQPHSSYVYSGLSGQPHFFYNR 1139
QY 1140 -----AAPPATGSSILL 1151
Db 1140 FPAPYLTPRYPPGSGHLSRVSVKTSILPRDQNGPLPCDSGFNKQROAAVAPATGSSILL 1199
QY 1152 SMTVDVVECKLROTEGLDQNMPOYCTTIKANINGRVLSCNIDELKEMWANNFGDW 1211
Db 1200 SMTVDVVECKLROTEGLDQNMPOYCTTIKANINGRVLSCNIDELKEMWANNFGDW 1259
QY 1212 LFRSMVLEMRSVESQVDPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTLN 1271
Db 1260 LFRSMVLEMRSVESQVDPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTLN 1319
QY 1272 FSFEELNTLGLDEGAPRHNLNWSQOTRTPSLSSINQDSSIEISKLTDKQVQAEYRDAY 1331
Db 1320 FSFEELNTLGLDEGAPRHNLNWSQOTRTPSLSSINQDSSIEISKLTDKQVQAEYRDAY 1379
QY 1332 REVIAQMSOLEGTSSTISGRSSPHSTYIICQSSSGGSIHSTLQERKGEKELQEDGR 1391
Db 1380 REVIAQMSOLEGTSSTISGRSSPHSTYIICQSSSGGSIHSTLQERKGEKELQEDGR 1439
QY 1392 KGFMLKRGDVIDYSSSGSVTNEASPLDPTTEDEKSDQSGSKLPGKSSRPSLFTQDL 1451

Db 1440 KSFLMKRGDVIDYSSSGSVTNEASPLDPTTEDEKSDQSGSKLPGKSSRPSLFTQDL 1499
QY 1452 KLKGGGLRYQKLPDEDESGTRGVQIIPHCSCMTIRTKELKAK-ORECASPOEHSAPERT 1510
Db 1500 KLKGGGLRYQKLPDEDESGTRGVQIIPHCSCMTIRTKELKAK-ORECASPOEHSAPERT 1557
QY 1511 FIKAKEVLSDALDKOSSDSGVRNNESSPNHSLHNEAADSQLEKANLIELEDEGHSGK 1570
Db 1558 FIKAKEVLSDALDKOSSDSGVRNNESSPNHSLHNEAADSQLEKANLIELEDEGHSGK 1617
QY 1571 RGMPSHLSGLQDPIIARMSICSEDKSPSECSLIASSPESWPAQKAYINLRTPSTVTL 1630
Db 1618 RGMPSHLSGLQDPIIARMSICSEDKSPSECSLIASSPESWPAQKAYINLRTPSTVTL 1677
QY 1631 NNNTAPTNRANQNFDEIEGIRETSQVILRPGSPNPTAVQENLKSMAHKSQSSSYTRL 1690
Db 1678 NNNTAPTNRANQNFDEIEGIRETSQVILRPGSPNPTAVQENLKSMAHKSQSSSYTRL 1737
QY 1691 SKDASELHAASESTGFGEERESIL 1715
Db 1738 SKDASELHAASESTGFGEERESIL 1762

RESULT 4
US-10-021-571-4
; Sequence 4, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTIB
; FILE REFERENCE: CHA011A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-571-4

Query Match 93.4%; Score 8266; DB 14; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
QY 1 MSVLIISQSVINYVEENIPALKALLEKCKDVDERNECGQTPLMAAEQGNVEIVKELKN 60
Db 1 MSVLIISQSVINYVEENIPALKALLEKCKDVDERNECGQTPLMAAEQGNLEIVKELKN 60
QY 61 GANCNLEDLNDWTALISASKEGHHIVELLKSGASLEHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLNDWTALISASKEGHHIVELLKCGVNLHRDMGWTALMWACYKGRDVV 120
QY 121 ELLLSHGANSVTLQVSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPPLVWAAR 180
Db 121 ELLLSHGANSVTLQVSVYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPPLVWAAR 180
QY 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTSQVKEILKNPNVNLTKDGN 240
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTSQVKEILKNPNVNLTKDGN 240
QY 241 LMTASKEGHIEIVQDILLDAGTYVNIIDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
Db 241 LMTASKEGHIEIVQDILLDAGTYVNIIDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
QY 301 QDNKNTALYWAKEGNATWVRDILQCNPDTEICTKDGSTPLIKATKMRNIEVVELLDKGA 360
Db 301 QDNKNTALYWAKEGNATWVRDILQCNPDTEICTKDGSTPLIKATKMRNIEVVELLDKGA 360

[illegible]

Qy	1441	SERP	SLFQTD	KLKGG	RLRYOK	PLSP	DEDESG	TRVQ	ITPHCS	QWIR	TKRLKAKQ	RECASP	1500
Db	1441	SER	SLFQTD	KLKGG	SLRYOK	PLSP	DEDESG	TRVQ	ITPHCS	DEDESG	TEESONT	-LLKXDDKDRKAE	GKVERVPKS 1499
Qy	1501	QEHSA	EPIRTF	IKAK	EYLS	DALLD	KKDSSD	SGVRSN	ESSPNH	SLHNEA	ADDSQ	LEKANLI	1560
Db	1500	PEHSA	EPIRTF	IKAK	EYLS	DALLD	KKDSSD	SGVRSN	ESSPNH	SLHNEA	ADDSQ	LEKANLI	1559
Qy	1561	ELEDE	HSGKRG	KMPH	SLSG	LQDPII	ARMSC	ISDCK	KSPSC	SLSIASS	PEESW	PACQKAYN	1620
Db	1560	ELED	SHSGKRG	KMPH	SLSG	LQDPII	ARMSC	ISDCK	KSPSC	SLSIASS	PEENW	PACQKAYN	1619
Qy	1621	LNRTP	SVTV	LNNAT	PNTNR	ANQNF	DEIEG	IRETSQ	IVILRP	GPSPNP	TAVONEN	LKSMHAK	1680
Db	1620	LNRTP	SVTV	LNNAS	PAFNR	ANQNF	DEMEG	IRETSQ	IVILRP	SSSPNP	TTIONEN	LKSMTHK	1679
Qy	1681	RSORSS	YTRLSK	DAASELH	-AASSE	STGFG	EERESIL	1715					
Db	1680	RSORSS	YTRLSK	DPPELH	AAASSE	STGFG	EERESIL	1715					

RESULT 5

US-10-117-229-2

; Sequence 2, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-2

Query Match

Best Local Similarity

Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

Qy	1	MSVL	ISQSV	INYVEE	ENIP	ALKAL	KKCKD	VD	VERNE	CGQT	PLMLAA	EQGN	VEIVK	ELIKN 60																																								
Db	1	MSVL	ISQSV	INYVEE	ENIP	ALKAL	KKCKD	VD	VERNE	CGQT	PLMLAA	EQGN	LEIVK	ELIKN 60																																								
Qy	61	GANC	NLED	LDN	WTAL	SASKE	GHHI	VEB	LLKSG	ASLEH	RD	MG	GTW	ALMWAC	YKGR	TDVV 120																																						
Db	61	GANC	NLED	LDN	WTAL	SASKE	GHVH	VEE	LLKCGV	NLEH	RD	MG	GTW	ALMWAC	YKGR	TDVV 120																																						
Qy	121	ELL	SHGAN	PSV	TG	LV	SVY	PII	WAAG	RHAD	I	VH	LLQNG	AKVNC	SDKY	GT	TPLV	YWAAR 180																																				
Db	121	ELL	SHGAN	PSV	TG	LQ	SV	PII	WAAG	RHAD	I	VH	LLQNG	AKVNC	SDKY	GT	TPLV	YWAAR 180																																				
Qy	181	KH	LEC	VGH	L	LA	MG	AD	VD	QEG	AN	SMT	AL	IV	AVK	GGY	T	QSVKE	IL	KEN	P	N	V	N	L	T	DK	GN	T	240																								
Db	181	KH	LEC	VGH	L	LA	MG	AD	VD	QEG	AN	SMT	AL	IV	AVK	GGY	T	QSVKE	IL	KEN	P	N	V	N	L	T	DK	GN	T	240																								
Qy	241	LMT	ASKE	G	H	TE	I	V	QD	L	L	D	AG	T	Y	N	I	P	D	R	S	G	D	T	V	L	I	G	A	V	R	G	G	H	VE	I	V	R	L	L	Q	K	YAD	I	D	I	R	G 300						
Db	241	LMT	ASKE	G	H	TE	I	V	QD	L	L	D	AG	T	Y	N	I	P	D	R	S	G	D	T	V	L	I	G	A	V	R	G	G	H	VE	I	V	R	L	L	Q	K	YAD	I	D	I	R	G 300						
Qy	301	QDN	K	T	AL	Y	WA	V	E	K	GN	AT	M	VR	D	I	L	Q	C	N	P	D	T	E	I	C	T	K	D	G	E	T	P	L	I	K	A	T	K	R	N	E	I	V	E	L	L	L	D	K	G	A 360		
Db	301	QDN	K	T	AL	Y	WA	V	E	K	GN	AT	M	VR	D	I	L	Q	C	N	P	D	T	E	I	C	T	K	D	G	E	T	P	L	I	K	A	T	K	R	N	E	I	V	E	L	L	L	D	K	G	A 360		
Qy	361	KV	SA	V	D	K	K	G	D	T	P	L	H	A	I	R	G	S	R	L	A	E	L	L	E	N	P	K	D	G	R	L	L	Y	R	P	N	K	A	G	E	T	P	N	I	D	C	S	H	O	K	S	I	L 420
Db	361	KV	SA	V	D	K	K																																															

Db 421 TQIFGARHLSPTEGDMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSFKLLKL 480
 Qy 481 EDEMKTFAGQOTBPLFQFQSWLIIVFLTLCLCGGLVPAFPVDTNLAIASLSFLALIYIF 540
 Db 481 EDEMKTFAGQOIIBPLFQFQSWLIIVFLTLCLCGGLVPAFTVHNPGLIAVSLSLALLIYIF 540
 Qy 541 FIVIFGGRREGESWNWAWALSTRLARHIGLYELLKPLMFVNPPELPQOTTKALPVRFLE 600
 Db 541 FIVIFGGRREGESWNWAWLSTLARHIGLYELLKPLMFVNPPELPQOTTKALPVRFLE 600
 Qy 601 TDYRLSSVGGTSLAEMIATLSDACERBEFGLATLFRVFRTEESQKKWKKTCCCLPS 660
 Db 601 TDYRLSSVGGTSLAEMIATLSDACERBEFGLATLFRVFKTEDTQKKWKKTCCCLPS 660
 Qy 661 FVIFLFTVGCIIAGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVNLNCRNTWQVLDLSLL 720
 Db 661 FVIFLFTIIGCIISGITLLAIFRVDPKHLTVNAVILISIASVVGGLAFVNLNCRNTWQVLDLSLL 720
 Qy 721 NSQKRLHSAASKLHKLSGSGMKVLKCEVELMARMAKTIDSTQNTQRLVLIIDGLDAC 780
 Db 721 NSQKRLHNAASKLHKLSGSGMKVLKCEVELMARMAKTIDSTQNTQRLVLIIDGLDAC 780
 Qy 781 EQDKVQLMDLTVRVLFSKGFPIAFASDPHIIKAINONLNSVLRDSNINNGHDMYRNIHV 840
 Db 781 EQDKVQLMDLTVRVLFSKGFPIAFASDPHIIKAINONLNSVLRDSNINNGHDMYRNIHV 840
 Qy 841 LPVFLNSRGLSNARKFVLTVSATNGDITCSDTTGTQEDTDRRVSONSIGEMTKLGSKTALN 900
 Db 841 LPVFLNSRGLSNARKFVLTVSATNGDVPDSDTGTQEDADRRVSONSIGEMTKLGSKTALN 900
 Qy 901 RRDYRBRQMTTROMSDPLTKLTVEDWFDSDISQPMRRLNIVSVTGRLLRANOIT 960
 Db 901 RRDYRBRQMTTROMSDPLTKLTVEDWFDSDISQPMRRLNIVSVTGRLLRANOIS 960
 Qy 961 FNWDLASWNLTEQWYRTSWLIYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 Db 961 FNWDLASWNLTEQWYRTSWLIYLEETEGIPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 Qy 1021 EIDGDIRNFVFLSSRTPVILVARDVKTFPLCTVNLDPKLEIITADVRAAREQINIGGLAY 1080
 Db 1021 EIDGDIRNFVFLSSRTPVILVARDVKTFPLCTVNLDPKLEIITADVRAAREQISIGGLAY 1080
 Qy 1081 PPLPLHEGPPRPSGYSQPSVCSASFPNGPFGVVSPPHSSYVYSGLSGPQHPFNRA 1140
 Db 1081 PPLPLHEGPPRPSGYSQPSVCSSTFNGPFGVVSPPHSSYVYSGMTGPQHPFNRG 1140
 Qy 1141 AVPATGSLILSSMTVDVCEKFLQIEGLDQNMMPQYCTTIKKANINGRVLSCNIDELK 1200
 Db 1141 SGPAFGPVLNLSLVDAVCEKFLQIEGLDQMLPQYCTTIKKANINGRVLACNIDELK 1200
 Qy 1201 KEMAMNFGDWHLPFRSMVLEKRSVSVVPPDPRFLNENSNAPVPHGESARRSHTELPLT 1260
 Db 1201 KEMAMNFGDWHLPFRSTVLEKRNASHVVPDPRFLSESSGPPAPHPGPARASHNELPHT 1260
 Qy 1261 ELSSQTPYTLNFSFEELNTGLDGAPRHSNLSQWQOTRTPPSLSSNSQDSSSIEISKLT 1320
 Db 1261 ELSSQTPYTLNFSFEELNTGLDGAPRHSNLSQWQOTRTPPSLSSNSQDSSSIEISKLT 1320
 Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGGSIHSTLEQERG 1380
 Db 1321 DKVQAEYRDAYREYIAQMSQLEGGFGSTISGRSSPHSTYYIMQSSGGSIHSLNEQEGK 1380
 Qy 1381 KEGELKQEDGRKSLFKRGDVIDYSSGVSVNEASPLDPITEEDEKSDQSGSKLLPGKKS 1440
 Db 1381 KDSFPKPDGRKSLFKRGDVIDYSSGVSVNDASPLDPITEEDEKSDQSGSKLLPGKKS 1440
 Qy 1441 SERSLFOTDLKLGKGLRYOKLPSDESDSGTGRVQITPHCSKMRITRKLKAKQREKASP 1500
 Db 1441 SERSLFOTDLKLGKGLRYOKLPSDESDSGTGRVQITPHCSKMRITRKLKAKQREKASP 1500
 Qy 1501 QEHSAEPRTTIFKAYEYLSADALLDKQSSDSGVSRSNESPNSHLSHNEAADSQLEKANLI 1560
 Db 1500 PEHSAEPRTTIFKAYEYLSADALLDKQSSDSGVSRSNESPNSHLSHNEAADSQLEKANLI 1559

RESULT 6

US-10-473-574-26
 ; Sequence 26, Application US/10473574
 ; Publication No. US20040116670A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;
 ; APPLICANT: TANG, Y. Tom; YUE, Henry;
 ; APPLICANT: KHAN, Farrah A.; ISON, Craig H.;
 ; APPLICANT: BAUGHN, Mariah R.; WARREN, Bridget A.;
 ; APPLICANT: DUGGAN, Brendan M.; THANGAVELU, Kavitha;
 ; APPLICANT: HONCHELL, Cynthia D.; AZIMZAI, Yalda;
 ; APPLICANT: ELLIOTT, Vicki S.; BURFORD, Neil;
 ; APPLICANT: DING, Li; YUE, Huibin;
 ; APPLICANT: BECHA, Shanya; EMERLING, Brooke M.;
 ; APPLICANT: RICHARDSON, Thomas W.; LEE, Soo Yeun;
 ; APPLICANT: BANDMAN, Olga; LAL, Preeti G.;
 ; APPLICANT: LEE, Sally; GIETZEN, Kimberly J.;
 ; APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
 ; APPLICANT: LEE, Ernestine A.; SWARNAKAR, Anita;
 ; APPLICANT: RING, Huijun Z.; JONES, Karen Anne
 ; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
 ; FILE REFERENCE: PF-0918 USN
 ; CURRENT APPLICATION NUMBER: US/10/473,574
 ; CURRENT FILING DATE: 2003-09-29
 ; PRIOR APPLICATION NUMBER: PCT/US02/09288
 ; PRIOR FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER: US 60/294,451
 ; PRIOR FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: US 60/291,870
 ; PRIOR FILING DATE: 2001-05-18
 ; PRIOR APPLICATION NUMBER: US 60/290,518
 ; PRIOR FILING DATE: 2001-05-10
 ; PRIOR APPLICATION NUMBER: US 60/288,609
 ; PRIOR FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/283,769
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: US 60/281,323
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/280,508
 ; PRIOR FILING DATE: 2001-03-29
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 26
 ; LENGTH: 1715
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 2755454CD1
 US-10-473-574-26

Query Match 93.4%; Score 8266; DB 16; Length 1715;
 Best local Similarity 92.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
 Qy 1 MSVLISQSVNYVEENIPALKLEKCKVDNERNCGQTPMLAARQGNVIVKELKN 60
 Db 1 MSVLISQSVNYVEENIPALKLEKCKVDNERNCGQTPMLAARQGNVIVKELKN 60


```
Qy 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Db 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Qy 121 ELLLSHGANSVTLQOYSVYPIIWAAGRHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANSVTLQOYSVYPIIWAAGRHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Qy 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTOSVKEILLKRNPNVNLTKDGNNTA 240
Db 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTOSVKEILLKRNPNVNLTKDGNNTA 240
Qy 241 LMIASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Db 241 LMIASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Qy 301 QDNKTALYVAKEGNATVWDILQCNPTETICTKDGTEPLIKATMRNIEVEELLKDKGA 360
Db 301 QDNKTALYVAKEGNATVWDILQCNPTETICTKDGTEPLIKATMRNIEVEELLKDKGA 360
Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRPNKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
Db 361 KVSADVKKGDTPLHVAIRGRSRLAELLRPNKDGRLLYRPNKAGETPYNIDCSHQKSIL 420
Qy 421 TQIFGARHLSTETDGMGLYSSALADILSPTMQPPICVGLYAGWGSKFLKKL 480
Db 421 TQIFGARHLSTETDGMGLYSSALADILSPTMQPPICVGLYAGWGSKFLKKL 480
Qy 481 EDEKMTAGQOTEPLFQPSWILVFLTLCCGLGLVAFVVDVNLATISFLALYIF 540
Db 481 EDEKMTAGQOTEPLFQPSWILVFLTLCCGLGLVAFVVDVNLATISFLALYIF 540
Qy 541 FIVYFGRRGESWNAWALSTLARIHIGYLELLFKLMFVNPPPELQTTKALPVRFPLF 600
Db 541 FIVYFGRRGESWNAWALSTLARIHIGYLELLFKLMFVNPPPELQTTKALPVRFPLF 600
Qy 601 TDYNRLSVGETSLAEMIATLSACEREFGLATRLFRVFRTEESQKKWKTCCLPS 660
Db 601 TDYNRLSVGETSLAEMIATLSACEREFGLATRLFRVFRTEESQKKWKTCCLPS 660
Qy 661 FVIFLFIIGCIAGTLLAIFRVPDKHLVTNAILISTASVVGAFVLCNRTWQVLSLL 720
Db 661 FVIFLFIIGCIAGTLLAIFRVPDKHLVTNAILISTASVVGAFVLCNRTWQVLSLL 720
Qy 721 NSQKRLHSAASKLHLKSGEFMKVLCVELMARMAKTIDSFQONQTRLVVIDGLDAC 780
Db 721 NSQKRLHSAASKLHLKSGEFMKVLCVELMARMAKTIDSFQONQTRLVVIDGLDAC 780
Qy 781 EODKVLQMLDTRVFLFSKGPPIAFASDPHIIKAINQNLNSVLRDSNNGHDYMRNIVH 840
Db 781 EODKVLQMLDTRVFLFSKGPPIAFASDPHIIKAINQNLNSVLRDSNNGHDYMRNIVH 840
Qy 841 LPVFLNRLSGLSNARKFLVTSATNGDITCSDTTQEDTDRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNRLSGLSNARKFLVTSATNGDITCSDTTQEDTDRVSONSLGEMTKLGSKTALN 900
Qy 901 RRDYRRQMTTITROMSPDLTKLLVTEDEWFSDI SPOTMRLNINVSVTGLLRANQIT 960
Db 901 RRDYRRQMTTITROMSPDLTKLLVTEDEWFSDI SPOTMRLNINVSVTGLLRANQIT 960
Qy 961 FNWDLASWINLTQWVPYRTSWLILYEETEGIPDMQTLKTIYERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWINLTQWVPYRTSWLILYEETEGIPDMQTLKTIYERISKNIPTTKDVEPLL 1020
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLRIBIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLRIBIADVRAAREQINIGGLAY 1080
Qy 1081 PPLPLEGPPPPPSGYSPASVCSASFNPGFPGVWSPOPHSSVYSGLSGPQPPFYNRA 1140
Db 1081 PPLPLEGPPPPPSGYSPASVCSASFNPGFPGVWSPOPHSSVYSGLSGPQPPFYNRA 1140
Qy 1141 AVPATGSSLLLSMTVDVCEKLRQIEGLDQNMMPQYCTTTIKKANINGRVLSCQNDLKL 1200
```

```
Db 1141 SGPAPOPVLNLSNVDVAVCEKLRQIEGLDQNMMPQYCTTTIKKANINGRVLSCQNDLKL 1200
Qy 1201 KEMAMNFGDWHLPFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGEGARSSHTPLT 1260
Db 1201 KEMAMNFGDWHLPFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGEGARSSHTPLT 1260
Qy 1261 ELSQTPYTLNFSFEELNTLGLDEGAPRHSNLWSQOTRRTPSSLNSQDSSIBISKLT 1320
Db 1261 ELSQTPYTLNFSFEELNTLGLDEGAPRHSNLWSQOTRRTPSSLNSQDSSIBISKLT 1320
Qy 1321 DKVOAERYDAYREYIAQMSQLEGGTSGSTISGRSPSHSTYYIIGQSSSGGSHSTLEQERG 1380
Db 1321 DKVOAERYDAYREYIAQMSQLEGGTSGSTISGRSPSHSTYYIIGQSSSGGSHSTLEQERG 1380
Qy 1381 KEGELKQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDPIITEEBEKSDQSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDPIITEEBEKSDQSGSKLLPGKKS 1440
Qy 1441 SERPSLFTQDLKLGGLRYOKLPFDEDESGTGRVQITPHCSKMITRKLKAKQRECA SP 1500
Db 1441 SERPSLFTQDLKLGGLRYOKLPFDEDESGTGRVQITPHCSKMITRKLKAKQRECA SP 1500
Qy 1501 QEHSABIPRIFIKAKYLSDALDKOSSDSGVSRNENSSPHSLHNEAADDQSOLEKANLI 1560
Db 1501 QEHSABIPRIFIKAKYLSDALDKOSSDSGVSRNENSSPHSLHNEAADDQSOLEKANLI 1560
Qy 1561 ELEDGHSGRKGMPSHLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYN 1620
Db 1561 ELEDGHSGRKGMPSHLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYN 1620
Qy 1621 LNRTPTSVTLNNAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHK 1680
Db 1621 LNRTPTSVTLNNAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHK 1680
Qy 1681 RSQRSSYTRLSKQASELH-AASSESTGFGERESIL 1715
Db 1681 RSQRSSYTRLSKQASELH-AASSESTGFGERESIL 1715
RESULT 7
US-10-117-229-4
; Sequence 4, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220PC
; FILE REFERENCE: 9U 104 RI
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-4
Query Match 92.8%; Score 8212; DB 14; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;
```

```
Qy 1 MSVLISQSVINLYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVVEIKELKN 60
Db 1 MSVLISQSVINLYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVVEIKELKN 60
Qy 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Db 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Qy 121 ELLLSHGANSVTLQOYSVYPIIWAAGRHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANSVTLQOYSVYPIIWAAGRHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
Qy 121 ELLLSHGANSVTLGL-YSVYPIIWAAGRHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 179
```

181 KGHLECVKHLAMGADVQDQGANSMALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
180 KGHLECVKHLAMGADVQDQGANSMALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 239
241 LMIASKEGHEIEIVQDLADAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
240 LMIASKEGHEIEIVQDLADAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 299
301 QDNKNTALYWAVERKGNATWDRDILQCNPDTEICTKDGETHPLIKATKMRNIEVVELLDKGA 360
300 QDNKNTALYWAVERKGNATWDRDILQCNPDTEICTKDGETHPLIKATKMRNIEVVELLDKGA 359
361 KVSADVKKGTTPHVAIRGRSRLAELLRNPKDGLLYRPNKAGETPNYINDCSHOKSIL 420
360 KVSADVKKGTTPHVAIRGRSRLAELLRNPKDGLLYRPNKAGETPNYINDCSHOKSIL 419
421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSFGKSLKKL 480
420 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSFGKSLKKL 479
481 EDEMKTFAGQOTPLFOFQSWLIVFLTLILCGGLGLVAFPVDTNLAIAISLSFLAIYIF 540
480 EDEMKTFAGQOTPLFOFQSWLIVFLTLILCGGLGLVAFVTHPNLGLIAVLSFLAIYIF 539
541 FIVYFGRREGSNNWAWLSRLARHIGVLELLPKLMFVNPPPELQOTTKALPVRFLE 600
540 FIVYFGRREGSNNWAWLSRLARHIGVLELLPKLMFVNPPPELQOTTKALPVRFLE 599
601 TDYRLSSVSGETSLAEMIAATLSDACEREFGLATRLFRVFRTEESQKKWKTKCCLPS 660
600 TDYRLSSVSGETSLAEMIAATLSDACEREFGLATRLFRVFKETDQKKWKTKCCLPS 659
661 FVILFLFVIGCIAGITLLAIIFRVDPKLTVNAILISIASVVGGLAFVLNCRTWQVLDL 720
660 FVILFLFVIGCIAGITLLAIIFRVDPKLTVNAILISIASVVGGLAFVLNCRTWQVLDL 719
721 NSQRKRLHSAASKLHKLKSGFMKVLKCEVELMARMAKTIDSTFQNTQRLVLIIDGLDAC 780
720 NSQRKRLHSAASKLHKLKSGFMKVLKCEVELMARMAKTIDSTFQNTQRLVLIIDGLDAC 779
781 EQDKVLQMLDTRVFLSKGPFIIAFASDPHIIKAINQNLNSVLDSNNGHDMYRNIVH 840
780 EQDKVLQMLDTRVFLSKGPFIIAFASDPHIIKAINQNLNSVLDSNNGHDMYRNIVH 839
841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVVSQNSLGMTKLSKLTALN 900
840 LPVFLNSRGLSNARKFLVTSATNGDVPSCDSTTGTQEDADRRVVSQNSLGMTKLSKLTALN 899
901 RRDYRRRQORITTRQMSFDLTKLVTDWFDSDISPTQWRRLLNIVSVTGRLLRANQIT 960
900 RRDYRRRQORITTRQMSFDLTKLVTDWFDSDISPTQWRRLLNIVSVTGRLLRANQIS 959
961 FNDRLASWINLQEQPYRTSWILYLEETEGLPDQWTKMYERISKNIPTTKDVEPLL 1020
960 FNDRLASWINLQEQPYRTSWILYLEETEGIPDQWTKMYERISKNIPTTKDVEPLL 1019
1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIADVRAAREQINIGGLAY 1080
1020 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIADVRAAREQISIGGLAY 1079
1081 PPLPLEHGP RPSPGSQPVASVCSASFNGFPFGVVSPOPHSYSGLSGQHPFVN-- 1138
1080 PPLPLEHGP RPSPGSQPVASVCSASFNGFPFGVVSPOPHSYSGMTGPOHPFVNR 1139
1139 -----RAAVP 1143
1140 FFAPLYLTPRYPGSOHLISRPVSKTSLPRDQNGNLEVIKEDAAEGLSPTDSSRSGSP 1199
1144 ATGSSLLLSMTVDVCEKLRQIEGLDNMMPOYCTTIKKANINGRVLSCNIDELKKEM 1203
1200 APGFVLLNSLNDVACEKLRQIEGLDQSMLPQYCTTIKKANINGRVLACQNDIDELKKEM 1259

1204 AMNFGDWHLFRSMVLEMRSVESOVVEDPRFLNENSSAPVPHGESARRSSHTLPLTELS 1263
1260 NMNFGDWHLFRSTVLEMRNAESHVVEDPRFLSESSSGPHAGEPARASHNELPHTELS 1319
1264 SQTPTLNFSPFELNLTGLDEGAPRHSNLWSQOTRRTPTSLSLNSQDSSIEISKLTDKV 1323
1320 SQTPTLNFSPFELNLTGLDEGAPRHSNLWSQOTRRTPTSLSLNSQDSSIEISKLTDKV 1379
1324 QAEYRDAYREYTAQMSOLEGSGTSSITISGRSSPHSYTYICOSSSGSIIHSTLEORKEG 1383
1380 QAEYRDAYREYTAQMSOLEGSGTSSITISGRSSPHSYTYICOSSSGSIIHSTLEORKEG 1439
1384 ELKQEDGRKSFMLMKRGDVIDYSSGVSTNEASPLDITEDEKSDQSGSKLLPGKSSER 1443
1440 EKPDDGRKSFMLMKRGDVIDYSSGVSTNDASPLDITEDEKSDQSGSKLLPGKSSER 1499
1444 PSLFQTDLKLKGGGLYQKLPSEDESGTGRVOITPHCSKMITRKLKAKQRECAPQEH 1503
1500 SSLFQTDLKLKGGGLYQKLPSEDESGTGRVOITPHCSKMITRKLKAKQRECAPQEH 1558
1504 SAEPIRTFKAKAYLSDALDKKSDSGVRSNESPNSHSLHNEAADDQLEKANILEF 1563
1559 SAEPIRTFKAKAYLSDALDKKSDSGVRSNESPNSHSLHNEAADDQLEKANILEF 1618
1564 DEHSGKRGMPHSLGLQDPIIARMSICSEDKSPSECSLIASSPEESWPACOKAYNLNR 1623
1619 DSHSGKRGMPHSLGLQDPIIARMSICSEDKSPSECSLIASSPEESWPACOKAYNLNR 1678
1624 TSTVTNNNTAPTNRANQNFDEIGRETSQVILRPGSPNPPTAVONENLKMMAHRSQ 1683
1679 TSTVTNNNTAPTNRANQNFDEIGRETSQVILRPGSPNPPTAVONENLKMMAHRSQ 1738
1684 RSYTSLKXDASELH--AASSESTGFEERESIL 1715
1739 RSYTSLKXDASELH--AASSESTGFEERESIL 1771

RESULT 8
US-10-117-229-11
; Sequence 11, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-117-229-11

Query Match 92.8%; Score 8212; DB 14; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;

1 MSVLIQSQVINYVEENIPALKALEKCKDVERNECGQTPLMAAEQGNVEIVKELLKN 60
1 MSVLIQSQVINYVEENIPALKALEKCKDVERNECGQTPLMAAEQGNVEIVKELLKN 60
61 GANCNLELDNNTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMACYKGRDVT 120
61 GANCNLELDNNTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMACYKGRDVT 120
121 ELLLSHGANPSVTGLQYVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPVWAAR 180
121 ELLLSHGANPSVTGL--YSYVPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPVWAAR 179
181 KGHLECVKHLAMGADVQDQGANSMALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240

Db 180 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTDKGNNTA 239
Qy 241 LMIASKEGHEIETVQDLADAGTYVNIIPDRSGDVTLCVAVRGHVEIVRALLQKADIDIRG 300
Db 240 LMIASKEGHEIETVQDLADAGTYVNIIPDRSGDVTLCVAVRGHVEIVRALLQKADIDIRG 299
Qy 301 QNKATLYAVVEKGNATVVRDILQCNPDTEICTKGETPLIKATQORNIIEVVELLDKGA 360
Db 300 QNKATLYAVVEKGNATVVRDILQCNPDTEICTKGETPLIKATQORNIIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPHVAIRGRSRLAELLRNPKDGLAYRKNKAGETPNYDNCQKHSIL 420
Db 360 KVSADVKKGDTPHVAIRGRSRLAELLRNPKDGLAYRKNKAGETPNYDNCQKHSIL 419
Qy 421 TQIFGARHLSPTETDGMGLDYLYSSALADILSEPTMOPPICVGLYAQWGSKSFLLKXL 480
Db 420 TQIFGARHLSPTETDGMGLDYLYSSALADILSEPTMOPPICVGLYAQWGSKSFLLKXL 479
Qy 481 EDEMTKTFAGQTEPLFQFSWLIIVFTLLCGGLGLVFPAPVDVNIINAIISFLALIYIF 540
Db 480 EDEMTKTFAGQTEPLFQFSWLIIVFTLLCGGLGLVFPAPVDVNIINAIISFLALIYIF 539
Qy 541 FIVYFGRREGESNNWAWLSTRLARHIGYLELLPKLMFVNPPELPTQTTKALPVRFLF 600
Db 540 FIVYFGRREGESNNWAWLSTRLARHIGYLELLPKLMFVNPPELPTQTTKALPVRFLF 599
Qy 601 TDYNRLSSVGGETSIAEMIALSDACEREFGLFATRLFRVFRTEBSQKKWKKTCCPLPS 660
Db 600 TDYNRLSSVGGETSIAEMIALSDACEREFGLFATRLFRVFRTEBSQKKWKKTCCPLPS 659
Qy 661 FVIFLFIIGCIIGITLLAIFRVDPKHITVNAIISIASVVGGLAFVLCNCRWQVLDL 720
Db 660 FVIFLFIIGCIIGITLLAIFRVDPKHITVNAIISIASVVGGLAFVLCNCRWQVLDL 719
Qy 721 NSQRKRLHSAASKLHLKSEGFMYLKEVELMARMKTIIDSTFQNTQRLVVIDGLDAC 780
Db 720 NSQRKRLHSAASKLHLKSEGFMYLKEVELMARMKTIIDSTFQNTQRLVVIDGLDAC 779
Qy 781 BQDKVQMLDVTVRVLFSGPIAFIPASDPHIIKAINQNLNVLDSNINGHYWRNIVH 840
Db 780 BQDKVQMLDVTVRVLFSGPIAFIPASDPHIIKAINQNLNVLDSNINGHYWRNIVH 839
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVVSQNSLGEMTKLASKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVVSQNSLGEMTKLASKTALN 899
Qy 901 RRDYRRRQMTTITRQMSFDLTKLLVTEWFDSDISQPMRRLNIVSVTGRLLRANQIT 960
Db 900 RRDYRRRQMTTITRQMSFDLTKLLVTEWFDSDISQPMRRLNIVSVTGRLLRANQIS 959
Qy 961 FNDRLASWINLTQWMPYRTSWLILYLBETEGLPDQMTLKTMYERISQNIPTTKOVBELL 1020
Db 960 FNDRLASWINLTQWMPYRTSWLILYLBETEGLPDQMTLKTMYERISQNIPTTKOVBELL 1019
Qy 1021 EIDGDIRNFVFLSRTPVLVARDVKTPLCTVNLDPKREIIADVRAAREQINIGGLAY 1080
Db 1020 EIDGDIRNFVFLSRTPVLVARDVKTPLCTVNLDPKREIIADVRAAREQINIGGLAY 1079
Qy 1081 PPLPLHEGPPRPPSGYOPASVCSSASFNPGVSPQPHSSYGLSGQPFFNYN -- 1138
Db 1080 PPLPLHEGPPRPPSGYOPASVCSSASFNPGVSPQPHSSYGLSGQPFFNYNRP 1139
Qy 1139 -----RAVP 1143
Db 1140 FFAPLYTPRYPGGSQHLISRPVKTSLPRDONNGLEVIKEDAEGLSPDTSRSGSP 1199
Qy 1144 ATGSSLLSMTVDVCEKLRQIEGLDQNMMPQCTTIKKNINGRVLSCQNDILKEM 1203
Db 1200 APGPVLLNSLVNDAVCEKLRQIEGLDQNMMPQCTTIKKNINGRVLSCQNDILKEM 1259
Qy 1204 AMNFGDWHLFRSMVLENRSVESQVVPDPREFLNENASAPVPHGESARRSHTEPLTELS 1263
Db 1260 NMNFGDWHLFRSMVLENRSVESQVVPDPREFLNENASAPVPHGESARRSHTEPLTELS 1319

Qy 1264 SOTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSOTRRTPSLSSLSNDSQSSISKLTDKV 1323
Db 1320 SOTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSOTRRTPSLSSLSNDSQSSISKLTDKV 1379
Qy 1324 QAEYRDAYREYIAQWSQLEGGTSGSTISGRSSPHSTYYIGOSSGGSIHSTLEQERKREG 1383
Db 1380 QAEYRDAYREYIAQWSQLEGGTSGSTISGRSSPHSTYYIGOSSGGSIHSTLEQERKREG 1439
Qy 1384 ELKQEDGRKSFMLKRGDVIDYSSGSVSTNEASPLDPIITEDEKSDQSGSKLLPGKKSER 1443
Db 1440 EPKPDGGRKSFMLKRGDVIDYSSGSVSTNEASPLDPIITEDEKSDQSGSKLLPGKKSER 1499
Qy 1444 PSLFOTDLKLGKGLRYOKLPDSDESGTGRVQTPHCSKMIRTKRLKAKQREKASPOEH 1503
Db 1500 SSLFOTDLKLGKGLRYOKLPDSDESGTGRVQTPHCSKMIRTKRLKAKQREKASPOEH 1558
Qy 1504 SAEPIRTFIKAKYLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYNLNR 1563
Db 1559 SAEPIRTFIKAKYLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYNLNR 1618
Qy 1564 DEHSGRGMPSHLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYNLNR 1623
Db 1619 DSHSGRGMPSHLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYNLNR 1678
Qy 1624 TPTVTTLNNTAPTNRANQNFDETEGIRETSQVILRPGSPNPNTAVONENLKSMAHKSQ 1683
Db 1679 TPTVTTLNNTAPTNRANQNFDETEGIRETSQVILRPGSPNPNTAVONENLKSMAHKSQ 1738
Qy 1684 RSSYTRLSDASELH-AAASESTGFEERESIL 1715
Db 1739 RSSYTRLSDASELH-AAASESTGFEERESIL 1771

RESULT 9

US-10-117-229-3
; Sequence 3, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION: Technologies, Inc.
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; TYPE: PRT
; LENGTH: 1184
; ORGANISM: Homo sapiens
US-10-117-229-3

Query Match 64.7%; Score 5730.5; DB 14; Length 1184;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAAECGNVIEVKELKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAAECGNVIEVKELKN 60
Qy 61 GANCNLEDLONWTALISASKEGHIHVEELLKSGASLEHRDMGCGTALMWACYKGRDVT 120
Db 61 GANCNLEDLONWTALISASKEGHIHVEELLKSGASLEHRDMGCGTALMWACYKGRDVT 120
Qy 121 ELLLSHGANSVTGLQYSVYPIIWAAGRHADIVHLLQLONGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANSVTGLQYSVYPIIWAAGRHADIVHLLQLONGAKVNCSDKYGTTPLVWAAR 179
Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTDKGNNTA 240
Db 180 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTDKGNNTA 239
Qy 241 LMIASKEGHEIETVQDLADAGTYVNIIPDRSGDVTLCVAVRGHVEIVRALLQKADIDIRG 300

```

240 LMIASKEGHEIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
301 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKDGSETPLIKATKRNIEVVELLDKGA 360
300 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKDGSETPLIKATKRNIEVVELLDKGA 359
361 KVSADVKKGGTTPHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPPYNDICSHOKSIL 420
360 KVSADVKKGGTTPHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPPYNDICSHOKSIL 419
421 TQIFGARHLSPTEITDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 480
420 TQIFGARHLSPTEITDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 479
481 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVFAFPVDVNTLAIASLSFLALIYIF 540
480 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVFAFPVDVNTLAIASLSFLALIYIF 539
541 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 600
540 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 599
601 TDYNRLLSSVGGETSLSAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCLPS 660
600 TDYNRLLSSVGGETSLSAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCLPS 659
661 FVIFLPIVGCIIAGITLLAIFRVDPKHLTWNALISIASVVGFLAVLNCRTWQVLDLSLL 720
660 FVIFLPIVGCIIAGITLLAIFRVDPKHLTWNALISIASVVGFLAVLNCRTWQVLDLSLL 719
721 NSQKRLHSAASKHLKLSGEGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 780
720 NSQKRLHSAASKHLKLSGEGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 779
781 EQDKVLQMLDTRVRLFSKGPFFIAIFASDPHIIKAINQNLSVLRDSNINNGHYMRNVH 840
780 EQDKVLQMLDTRVRLFSKGPFFIAIFASDPHIIKAINQNLSVLRDSNINNGHYMRNVH 839
841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 899
901 RRDYRRRQMQRTITRQMSFDLTCLAVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960
900 RRDYRRRQMQRTITRQMSFDLTCLAVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 959
961 FNWDRLASWINLTQWQPYRTSWLILYLEETEGLPDQMTLKTMYERIISKNIPTTKDVEPLL 1020
960 FNWDRLASWINLTQWQPYRTSWLILYLEETEGIPDQMTLKTMYERIISKNIPTTKDVEPLL 1019
1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQINIGGLAY 1080
1020 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGGLAY 1079
1081 PPLPHGPPRPSPGYSQPPSCVCSASFNGPFGGVVSPPHSSYSGLSGPOHPFYNR 1139
1080 PPLPHGPPRPSPGYSQPPSCVCSSTSGNPGFAGGVVSPPHSSYSGMIGPOHPFYNR 1138

```

RESULT 10

```

US-10-117-229-9
; Sequence 9, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117, 229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 9

```

```

; LENGTH: 1184
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-117-229-9

Query Match      64.7%; Score 5730.5; DB 14; Length 1184;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

QY 1 MSVLIISQSVINYYEENIPALKALLEKCKDNDVERNECGOTPLMLAAEQGNVEIVKELKN 60
DB 1 MSVLIISQSVINYYEENIPALKALLEKCKDNDVERNECGOTPLMLAAEQGNLEIVKELKN 60
QY 61 GANCNLEDNDNTALISASKEGHHIVVEBELLKSGASLEHRDMGWTALMWACYKGRDVV 120
DB 61 GANCNLEDNDNTALISASKEGHHIVVEBELLKSGVLEHRDMGWTALMWACYKGRDVV 120
QY 121 ELLSHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPVWAAR 180
DB 121 ELLSHGANPSVTGL-YSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPVWAAR 179
QY 181 KGHLECVKHLLANGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGN 240
DB 180 KGHLECVKHLLANGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGN 239
QY 241 LMIASKEGHEIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
DB 240 LMIASKEGHEIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
QY 301 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKDGSETPLIKATKRNIEVVELLDKGA 360
DB 300 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKDGSETPLIKATKRNIEVVELLDKGA 359
QY 361 KVSADVKKGGTTPHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPPYNDICSHOKSIL 420
DB 360 KVSADVKKGGTTPHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPPYNDICSHOKSIL 419
QY 421 TQIFGARHLSPTEITDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 480
DB 420 TQIFGARHLSPTEITDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 479
QY 481 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVFAFPVDVNTLAIASLSFLALIYIF 540
DB 480 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVFAFPVDVNTLAIASLSFLALIYIF 539
QY 541 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 600
DB 540 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 599
QY 601 TDYNRLLSSVGGETSLSAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCLPS 660
DB 600 TDYNRLLSSVGGETSLSAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCLPS 659
QY 661 FVIFLPIVGCIIAGITLLAIFRVDPKHLTWNALISIASVVGFLAVLNCRTWQVLDLSLL 720
DB 660 FVIFLPIVGCIIAGITLLAIFRVDPKHLTWNALISIASVVGFLAVLNCRTWQVLDLSLL 719
QY 721 NSQKRLHSAASKHLKLSGEGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 780
DB 720 NSQKRLHSAASKHLKLSGEGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 779
QY 781 EQDKVLQMLDTRVRLFSKGPFFIAIFASDPHIIKAINQNLSVLRDSNINNGHYMRNVH 840
DB 780 EQDKVLQMLDTRVRLFSKGPFFIAIFASDPHIIKAINQNLSVLRDSNINNGHYMRNVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
DB 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 899
QY 901 RRDYRRRQMQRTITRQMSFDLTCLAVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIT 960
DB 900 RRDYRRRQMQRTITRQMSFDLTCLAVTEDWFSDISPQTMRRLLNIVSVTGRLLRANQIS 959

```

Qy 961 FNDRLASWNLTEQWPVRTSWLILYLEETECPLDQMTLKTMYERISKNIPPTTKDVEPLL 1020
Db 960 FNDRLASWNLTEQWPVRTSWLILYLEETEGIPDQMTLKTMYERISKNIPPTTKDVEPLL 1019
Qy 1021 EIDGDIRNFVFLSRTPLVARDVKPLCTVNLDPKRLRIIADVRAAREQINIGGLAY 1080
Db 1020 EIDGDIRNFVFLSRTPLVARDVKPLCTVNLDPKRLRIIADVRAAREQISIGGLAY 1079
Qy 1081 PPLPLHEGPPRPPSGYSQPASVCSASFNPGFPGVSPQPHSSYSGSLGQPQHPFYNR 1139
Db 1080 PPLPLHEGPPRPPSGYSQPPSCVSTSFNGFAGGVSPQPHSSYSGMTQHPFYNR 1138
RESULT 11
US-10-106-698-6378
; Sequence 6378, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6378
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (244)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (337)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6378
Query Match 33.5%; Score 2961.5; DB 14; Length 705;
Best Local Similarity 95.5%; Pred. No. 1e-195;
Matches 567; Conservative 12; Mismatches 14; Indels 1; Gaps 1;
Qy 62 ANCNLEDLDNWTALISASKEGHIHVEELLKSGASLEHRDGGWTALMWACYKGRDVE 121
Db 28 SNCNLEDLDNWTALISASKEGHVHVEELLKCGVNLHRDGGWTALMWACYKGRDVE 87
Qy 122 LLLSHGANPSVTGLQYSVYPIIWAAGRHDIVHLLQNGAKVNCSDKYGTTPPLVWAARK 181
Db 88 LLLSHGANPSVTGL-QYSVYPIIWAAGRHDIVHLLQNGAKVNCSDKYGTTPPLVWAARK 146
Qy 182 GHLECVKHLAMGADVDEGANSMTALIVAKGGYTSVKELKRNPNVNLTDKGNLTAL 241
Db 147 GHLECVKHLAMGADVDEGANSMTALIVAKGGYTSVKELKRNPNVNLTDKGNLTAL 206
Qy 242 MIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQ 301
Db 207 MIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQ 266
Qy 302 DNKTALYWAKEGNATVNRDILQCNPDTEICTKDGTEPLIKATKORNIIEVVELLDKGA 361
Db 267 DNKTALYWAKEGNATVNRDILQCNPDTEICTKDGTEPLIKATKORNIIEVVELLDKGA 326
Qy 362 VSAVDKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPNIDCSHQSILT 421
Db 327 VSAVDKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPNIDCSHQSILT 386
Qy 422 QIFGARHLSPTETDGMGLYSSALADILSEPTMQPPICVGLYAWGSGKSFLLKKLE 481

Db 387 QIFGARHLSPTETDGMGLYSSALADILSEPTMQPPICVGLYAWGSGKSFLLKKLE 446
Qy 482 DEMKTFAGQOETPLFQPSWLVFLTLCCGGLGVFAFPDNTNLAIALISFLALYIFF 541
Db 447 DEMKTFAGQOETPLFQPSWLVFLTLCCGGLGVFAFPDNTNLAIALISFLALYIFF 506
Qy 542 IVIYFGRRREGESNNWAWALSTRLARHIGYLELFLKLMFVNPPELPEQTTKALPVRFLT 601
Db 507 IVIYFGRRREGESNNWAWALSTRLARHIGYLELFLKLMFVNPPELPEQTTKALPVRFLT 566
Qy 602 DYNRLSSVGGTSLAEMIATLSDACEREFGLATRLFRVFTERSQGGKKWKKT 655
Db 567 DYNRLSSVGGTSLAEMIATLSDACEREFGLATRLFRVFTEDTQGGKKKNS 620
RESULT 12
US-09-835-788A-17
; Sequence 17, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 17
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-17
Query Match 31.4%; Score 2782.5; DB 9; Length 551;
Best Local Similarity 96.7%; Pred. No. 1.6e-183;
Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;
Qy 102 MGGTALMWACYKGRDVEVLLLSHGANSVVTGLQYSVYPIIWAAGRHDIVHLLQNG 161
Db 1 MGGTALMWACYKGRDVEVLLLSHGANSVVTGL-QYSVYPIIWAAGRHDIVHLLQNG 59
Qy 162 AKVNCSDKYGTTPPLVWAARKGHLECVKHLAMGADVDEGANSMTALIVAKGGYTSVK 221
Db 60 AKVNCSDKYGTTPPLVWAARKGHLECVKHLAMGADVDEGANSMTALIVAKGGYTSVK 119
Qy 222 EILKRNPNVNLTDKGNLTALMIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRG 281
Db 120 EILKRNPNVNLTDKGNLTALMIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRG 179
Qy 282 HVEIVRALLQKYADIDIRGQDNKTALYWAKEGNATVNRDILQCNPDTEICTKDGTEPLI 341
Db 180 HVEIVRALLQKYADIDIRGQDNKTALYWAKEGNATVNRDILQCNPDTEICTKDGTEPLI 239
Qy 342 KATKORNIIEVVELLDKGAQVSAVDKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYR 401
Db 240 KATKORNIIEVVELLDKGAQVSAVDKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYR 299
Qy 402 NKAGETPNIDCSHQSILTQIFGARHLSPTETDGMGLYSSALADILSEPTMQPPI 461
Db 300 NKAGETPNIDCSHQSILTQIFGARHLSPTETDGMGLYSSALADILSEPTMQPPI 359
Qy 462 CVGLYAWGSGKSFLLKKLEDEMKTTFAGQOETPLFQPSWLVFLTLCCGGLGVFAFPV 521
Db 360 CVGLYAWGSGKSFLLKKLEDEMKTTFAGQOETPLFQPSWLVFLTLCCGGLGVFAFTV 419

Qy 522 DTNLAIAISLFLALIIYFFIVYFGRRGESNNWAWALSTRLARHIGVLELLFKLMFV 581
Db 420 HPNLGIAVSLSFLALIIYFFIVYFGRRGESNNWAWLSTRLARHIGVLELLKLMFV 479
Qy 582 NPPELPQTTKALPVRFLEFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 641
Db 480 NPPELPQTTKALPVRFLEFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 539
Qy 642 RTEBSQKKKKWK 653
Db 540 KTEDTQKKKKK 551

RESULT 13

US-10-175-042-17
; Sequence 17, Application US/10175042
; Publication No. US200301810A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides,
; TITLE OF INVENTION: Polypeptides, and
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/10/175,042
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 09/835,788
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-175-042-17

Query Match 31.4%; Score 2782.5; DB 14; Length 551;
Best Local Similarity 96.7%; Pred. No. 1.6e-183;
Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;
Qy 102 MGGWTALMWACYKGRDVTVELLSHGANGPSVTGL-YSVYPIIWAAGRGHADI VHLILLQNG 59
Db 1 MGGWTALMWACYKGRDVTVELLSHGANGPSVTGL-YSVYPIIWAAGRGHADI VHLILLQNG 59
Qy 162 AKVNCSDKYCTTPLVWAARKHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTOSVK 221
Db 60 AKVNCSDKYCTTPLVWAARKHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTOSVK 119
Qy 222 EILKRNPNVNLTKDGNLTALMIASKEGHIIEIVQDLLDAGTYVNI PDRSGTTLVIGAVRG 281
Db 120 EILKRNPNVNLTKDGNLTALMIASKEGHIIEIVQDLLDAGTYVNI PDRSGTTLVIGAVRG 179
Qy 282 HVEIVRALLQYADIDIRGOONKTALYWAKEGNATVVRDI LQCNPDTEICTKDGTEPLI 341
Db 180 HVEIVRALLQYADIDIRGOONKTALYWAKEGNATVVRDI LQCNPDTEICTKDGTEPLI 239
Qy 342 KATKRNIEVVELLLDKGAVSADKKGDTPLHVAIRGRSRLAEILLRNPKDGRLLYRP 401
Db 240 KATKRNIEVVELLLDKGAVSADKKGDTPLHVAIRGRSRLAEILLRNPKDGRLLYRP 299
Qy 402 NKAGETPNYDCSHQSKSILTIQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPI 461
Db 300 NKAGETPNYDCSHQSKSILTIQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPI 359
Qy 462 CVGLIYAQWGSKGKSFLLKKLEDEMKTFAGQQTPELPQFSWLI VFLTLLCCGLGLGVFAFPV 521
Db 360 CVGLIYAQWGSKGKSFLLKKLEDEMKTFAGQQTPELPQFSWLI VFLTLLCCGLGLGVFAFPV 419

Qy 522 DTNLAIAISLFLALIIYFFIVYFGRRGESNNWAWALSTRLARHIGVLELLFKLMFV 581
Db 420 HPNLGIAVSLSFLALIIYFFIVYFGRRGESNNWAWLSTRLARHIGVLELLKLMFV 479
Qy 582 NPPELPQTTKALPVRFLEFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 641
Db 480 NPPELPQTTKALPVRFLEFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 539
Qy 642 RTEBSQKKKKWK 653
Db 540 KTEDTQKKKKK 551

RESULT 14

US-10-021-571-8
; Sequence 8, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Haeyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTIB
; FILE REFERENCE: CHAO11A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Drosophila
US-10-021-571-8

Query Match 26.8%; Score 2376.5; DB 14; Length 1498;
Best Local Similarity 37.6%; Pred. No. 1e-154;
Matches 552; Conservative 272; Mismatches 527; Indels 117; Gaps 36;
Qy 1 MSVLISQSVINYYVEENIPALKALLE-KCKVDVERNECGQTPPLMLAAEQGNVIVKELLK 59
Db 23 MGSGLGHRALLQYIDNNDISGLRAILDSRHLLTIDDRDENATTVLMVYVAGRLTAFVREFLA 82
Qy 60 NGANCNLELDLNNWTALISAKSGHHIIVEELLSKGASLEHRDMGGWTALMWACYKGRDVT 119
Db 83 RGADVQAEUDDNNWTALLCASRNGHLVDVQLLDHGAEEVHRDNGGWTSLMWAAYRGHTEL 142
Qy 120 VELLISHGANPSVTGLQYSYVYPIIWAAGRGHADI VHLILLQNGAKVNCSDKYCTTPLVWAA 179
Db 143 VRLLLDKGADGNAHG-NYHLGALLWAAGRYKDI VELLVQRGAKVNVGDKYGTITALVWAC 201
Qy 180 RKGHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTOSVKEILKRNPNVNLTKDGN 239
Db 202 RRGNVEIVDTLLKAGANVDTAGYSWTPLLVAAGGHTDCVSSILEKKNVNALDKGMT 261
Qy 240 ALMIASKEGHIIEIVQDLLDAGTYVNI PDRSGTTLVIGAVRGHVEIVRALLQYADIDIR 299
Db 262 ALCIASREGQDIAASLIAAGAYINIQDRGADTPLIHAVKAGHRTVVEALLKGHADVDAQ 321
Qy 300 GQDNKTALYWAKEGNATVVRDI LQCNPDTEICTKDGTEPLIKATKORNIEVVELLLDKG 359
Db 322 GKDRKTAIYAVEKGHTPIVKLLLATNPDLSEATKDGTPLLRAVRNRNLEIVHLLLDK 381
Qy 360 AKVSAVDKDGTPPLHVAIRGRSRLAEILLRNPKDGRLLYRPNKAGETPNYDCSHQSKI 419
Db 382 AKVTASDKRGDTCLHTAMRARSKTIIVEALLRNPKHSQLLYRANKAGETPNYDCSHQKTI 441
Qy 420 LTIQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPI CVGLIYAQWGSKGKSFLLKK 479
Db 442 LGOVFGARRINTNDESEGMLGHLYELISSALADVLSSEPTLTTPITVGLIYAKWGSKGKSFLLKK 501
Qy 480 LEDEMKTFAGQQTPELPQFSWLI VFLTLLCCGLGLGVFAFPV-DTNLAIAISLFLALIIY 538

Db 502 LRDENNFARQWABPIRTSGLLFTVCLVHVALIGTIVGLTSWAVGVSAAGFLLAY 561
Qy 539 IFFIIVYFGREGESMNWALSTRLARHIGYLELLFKLMFVNPPELPEQTICALPVRF 598
Db 562 LLLAAVRYCNYQ--MDMQWASVQGLEKRWLRLELIQVAFCHPPG-PQSDSQAKPVRF 618
Qy 599 LFTDYNRLSSVGGTSLAEMIATISDACERBFGFLATLRFVFRTE--ESQGGKKKKWTK 656
Db 619 HFAEANGASPT-GDCAVHMLAALLDATESHYGWLATLRYAFRPKCLKVDVGNWRMRMC 677
Qy 657 CLPSFVFLFVGCIIAGITLAP-----RVDPKHLTVNAILISIASVGLAFVLCRT 711
Db 678 CIPVILPELALVTVTGISTVAYFTFADEKEKEHILV--ALYVIAVMGTLICHTLHV 735
Qy 712 WQVLDLSLNQRKLHSAASKJLHKLSEGMVKLVKCEVELMARMAKTIDFTQNRVLV 771
Db 736 LAKVFSVSLFTSHIRVLEAV-----RSSESAPLTMGAEVAVMTDMVKCLDAFTNQSELV 791
Qy 772 VIIDGLDACEQDKVLQMLDVTVRVLF--KGPIAIFASDPHIIKAINQNLSNVLRSNI 829
Db 792 GVIDALDSCDTERILTLNAVQTLTSSPNRPVLLISVDPHVIAKAAEANSRRLFTTEGI 851
Qy 830 NGHDMYRNIVHLPVFLSNGLSNARKPLVTS-----ATNGDITCSD--TTGTQEDTDRV 882
Db 852 GGHDFLRNLHLPVYLQNSGLRKVQRAQMTALLPKRSGGQYQTDGPTLG-HSVSARRL 910
Qy 883 SQNS--LGEMTKL-----GSKTALNR-----RDYRRQWQRTITRQMSFDLTKL 925
Db 911 SNASEIISQEKLRGARGGGKGLRLSESVASSTGSLHLRLGQNPQTV-----LDLSRI 965
Qy 926 LVTEWFSDISPQWRRLNLTIVSUGRLLRANQITFNWDLASWINLQEPYRTSLMIL 985
Db 966 VLTDDYFSDVNPWRMLNMVYITVRLKAKAFIEFSWYRLSSINLQEPWPLRAIMVL 1025
Qy 986 YLEE--TEGLPDQMTLKWYERISKNPTTKVDEPLEIDGDIRNFVFLSSRTPVLAR 1043
Db 1026 HHDQFMSNADSVSLOSVEKJRPKLAYLREAAPLLELDRDRKLDLQFLHKSLLVA 1085
Qy 1044 DVKTFLPCTVNLDPKLRRIADVAAREQIINIGLAYPPLPHEGP-----PRPPSGYSQP 1099
Db 1086 DLRIFLPFTINLDPYLRKVLKEDQOTIE--DEGSLVIOARPSVNTWROFPAPTYVPSP 1143
Qy 1100 ASVCSSASFNGPFGVVPQPHSYSGLSGQPHFYNRAAVPATGSSLL---LSSMTV 1156
Db 1144 QAYPPYQFQNEYAPN---ELSRNLSTSTEPVTLNPS-DSFGDDILQTKLTLTV 1198
Qy 1157 DVCEKLRQIEGLQNMVPOCTTIKKANINGRVLSCNIDELKEWAMNFGDHLFRSM 1216
Db 1199 EGVISLDRIBDM-KPALPKLAPVLRENAINGRVLKHCMDPLKSVLGSFGHWELFRLL 1257
Qy 1217 VLEMRVSE-----SQVVPEDPRFLNENSSAPV-----PHGEGARRS---SHTELP 1258
Db 1258 ITTLRECLRPKQORQOQOQGALEAPSNVPMIKDVTDALMQPPRESLRKNSVSHMEKQ 1317
Qy 1259 LT---ELSSQPTTYLNTFSFELNLTGLDEGAPRHSNLSWQOTRTTPSLSL-----N 1308
Db 1318 VTLEQICGTLQTLNEAYE-DVASSRSPPTGEMLAQAOLQAPRESSEFGSPSDD 1376
Qy 1309 SQDSIEISKLTDKVQ---ABY-RDAYREYTAQMSL-----EGGTGSSTI---SGRSS 1355
Db 1377 QKQYGVKISNNNNNNQVLAENYNSVSHSLSQSLTLVAPVGVGGGGSHLNGNDL 1436
Qy 1356 PHSTYVYIGQSSS-----GGSIHSTLEQ 1377
Db 1437 SDSTLDLMDVSVFGGGGGGTHRASRQ 1464

RESULT 15
US-10-149-819-9
; Sequence 9, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1
US-10-149-819-9

Query Match 25.5%; Score 2260; DB 14; Length 513;
Best Local Similarity 86.6%; Pred. No. 2.1e-147;
Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;

Qy 1203 MAMNFGDHLFRSVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTEL 1262
Db 1 MAMNFGDHLFRSVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 60
Qy 1263 SSQTPYTLNFSFELNLTGLDEGAPRHSNLSWQOTRTTPSLNLSQDSSEIISKLTKD 1322
Db 61 SSQTPYTLNFSFELNLTGLDEGAPRHSNLSWQOTRTTPSLNLSQDSSEIISKLTKD 120
Qy 1323 VQAEYRDAYREYTAQMSQLEGGTSGSTISGRSSPHSTYIYGSSSGSIHSTLQERKGE 1382
Db 121 VQAEYRDAYREYTAQMSQLEGGPGSTTISGRSSPHSTYIYGQSSSGSIHNSLQERKGD 180
Qy 1383 GELKQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDPTTEDEKSDQSGSKLLPKKSSSE 1442
Db 181 SEPKDDGRKSFMLKRGDVIDYSSGVSTNDASPLDPTTEDEKSDQSGSKLLPKKSSSE 240
Qy 1443 RPSLFQTDLKLKGGGLRYQKLPSEDESGTGRVQITPHCSMIRTKRLKAKQRECAPQOE 1502
Db 241 RSSLFQTDLKLKGGGLRYQKLPSEDESGTGRVQITPHCSMIRTKRLKAKQRECAPQOE 299
Qy 1503 HSAEPIRTFKAKYELSDALLDKKSDSGVRSNENSSPHSLHNEAADDSSOLEKANLTEL 1562
Db 300 HSAEPIRTFKAKYELSDALLDKKSDSGVRSNENSSPHSLHNEAADDSSOLEKANLTEL 359
Qy 1563 EDEGHSKGRGPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPESWAPACQAYNLN 1622
Db 360 EDDSHSGKRGPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPESWAPACQAYNLN 419
Qy 1623 RTSTVTTLNNTAPTNRANQNFDBIEGIRETSQVILRPGSPNPNTAVQENLKNMAHKRS 1682
Db 420 RTSTVTTLNNSAPNANQNFDEMEGIRETSQVILRPGSPNPNTAVQENLKNMAHKRS 479
Qy 1683 QRSSYTLKSDASELH--AASSESTGFGCEERESIL 1715
Db 480 QRSSYTLKSDPPPELHAAASSESTGFGCEERESIL 513

Search completed: February 5, 2005, 19:54:49
Job time : 156 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:16:39 ; Search time 31 Seconds
(without alignments)
5322.960 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSESTGGBERESIL 1715
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5073.5	57.3	1031	2 T43458	hypothetical prote
2	1848.5	20.9	1398	2 T21884	hypothetical prote
3	1024	11.6	246	2 T46392	hypothetical prote
4	556	6.3	1881	1 SUHUK	ankyrin 1, erythro
5	555	6.3	426	2 AE2149	hypothetical prote
6	553	6.2	1856	2 B35049	ankyrin 1, erythro
7	553	6.2	1880	2 A35049	ankyrin 1, erythro
8	552	6.2	1961	2 T42716	ankyrin 3, splice
9	551.5	6.2	1940	2 T42715	ankyrin 3, splice
10	551	6.2	1943	2 T42713	ankyrin 3, splice
11	550	6.2	3924	2 S37431	ankyrin 2, neurona
12	527.5	6.0	1862	2 I49502	ankyrin - mouse
13	526	5.9	4377	2 A55575	ankyrin 3, long sp
14	516.5	5.8	1848	2 S37771	ankyrin, erythrocy
15	510.5	5.8	1765	2 T42714	ankyrin 3, splice
16	479	5.4	2039	2 T15347	ankyrin-related un
17	464.5	5.2	1549	2 T13940	ankyrin - fruit fl
18	429	4.8	2584	2 T24158	hypothetical prote
19	429	4.8	2606	2 T24157	hypothetical prote
20	414.5	4.7	1423	1 T37275	death-associated p
21	409.5	4.6	791	2 T42691	hypothetical prote
22	397	4.5	1062	2 T30255	inversin - mouse
23	384.5	4.3	557	2 T46507	hypothetical prote
24	382	4.3	1062	2 T14151	Inv protein - mous
25	378.5	4.3	323	2 B47169	ankyrin-like repea
26	361	4.1	397	2 T46445	hypothetical prote
27	356	4.0	679	2 B45771	2-5A-dependent RNA
28	347.5	3.9	741	2 A45771	2-5A-dependent RNA
29	338.5	3.8	991	2 T25412	hypothetical prote

30	335	3.8	1401	2 S11527	alpha-latrotoxin p
31	325	3.7	1188	2 T19552	hypothetical prote
32	322.5	3.6	441	2 S74626	erythroid ankyrin
33	321.5	3.6	934	1 H71274	probable ankyrin -
34	302	3.4	1058	2 D82654	ankyrin-like prote
35	301	3.4	2004	2 D88948	ankyrin-like prote
36	300.5	3.4	1411	2 S30355	alpha-latrotoxin
37	296.5	3.3	627	2 E86190	hypothetical prote
38	294	3.3	1184	2 T00253	gene Ankhzn protei
39	292	3.3	606	2 AC2508	hypothetical prote
40	288	3.3	1083	2 S48460	probable membrane
41	287	3.2	517	2 T48283	ankyrin-like prote
42	286.5	3.2	662	2 E84725	ankyrin-like prote
43	285	3.2	1001	2 S30385	G9a protein - huma
44	279.5	3.2	368	2 T18184	ankyrin repeat pro
45	279	3.2	543	2 C86212	hypothetical prote

ALIGNMENTS

RESULT 1

T43458
hypothetical protein DKFZp434F0621.1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43458
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z22517
A:Accession: T43458
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1031 <AAA>
A:Cross-references: UNIPROT:Q9UF42; EMBL:AL133620
A:Experimental source: adult testis; clone DKFZp434F0621
C:Genetics:
A:Note: DKFZp434F0621.1

Query Match 57.3%; Score 5073.5; DB 2; Length 1031;
Best Local Similarity 95.2%; Pred. No. 8.1e-291;
Matches 974; Conservative 23; Mismatches 13; Indels 13; Gaps 1;

Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN	60
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN	60
Qy	61	GANCNLEDLNDWTALISASKEGHIHIVEELKSGASLEHRDMGGWTALMWACYKGRD	120
Db	61	GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALMWACYKGRD	120
Qy	121	ELLISHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLONGAKVNCSDKYGTTPLVWAAR	180
Db	121	ELLISHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLONGAKVNCSDKYGTTPLVWAAR	180
Qy	181	KGHLCECVKHLAMGADVDGANSMTALIVAVKGYTQSVKEILKRPNNVNLTKDGN	240
Db	181	KGHLCECVKHLAMGADVDGANSMTALIVAVKGYTQSVKEILKRPNNVNLTKDGN	240
Qy	241	LMTASKEGHEIIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQYADIDIRG	300
Db	241	LMTASKEGHEIIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQYADIDIRG	300
Qy	301	QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA	360
Db	301	QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA	360
Qy	361	KVSAVDKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRNKAGETPYNIDCSHQSIL	420
Db	361	KVSAVDKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRNKAGETPYNIDCSHQSIL	420
Qy	421	TQIFGARHLSPTDGDMLGYDLYSSALADILSPTMQPPTCVGLYAQWGSCKSFLLKKL	480


```
Db 154 QHENVVAHLINYGTKGVLPALHIAARNDDRTAAVLLQNDPNPDVLSKGTGTPHLIA 213
QY 84 -----IHIVEELKSGASLEHRDMGWMTALMAWACYGRDVTVELLSHGANPSVTLQVYS 138
Db 214 AHYENLVNVAQLLNRGASVNTFQNGITPLHIAARRGNVIMVRLLLDRGAQIE-TTKDE 272
QY 139 VYPIIWAAGRGAHIVHLLQNGAKNCSKYGTTPLVWAARKHLECVKHLAMGADV 198
Db 273 LTPHCAARNGHVRISEILLDHGAPQAKTKNGLSPIHMAAQGDHLDCKVRLLOYDAEID 332
QY 199 -----QEGA-----NSMTALIVAVKGYTQSVKEILK 225
Db 333 DITLDHLTPHVAACHGHRVAKVLLDKGAPNSRALNGFTPLHIAKKNHVRVMEILLK 392
QY 226 RNPVNLTDKGNVTALMIASKEGHIIVQDILLDAGTVVNIPIDRSGDTVLIGAVRGHVEI 285
Db 393 TGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSKVETPLHMAARAGHTEV 452
QY 286 VRALLQKYADIDIRGQDNKTALYAVEKGNATVDRDILQCNPDTEICTKGETPLIKATK 345
Db 453 AKYLLQNKAKVNAKADQDPLHCAARIHTNNMVKLLLENANPNLATTAGHTPLHIAAR 512
QY 346 MRNIEVVELLDGAKVSAVDKKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAG 405
Db 513 EGHVEITVALLKEASQACMTKGGFTPLHVAKYGVKRVVAELLE-RDA----HNAAG 566
QY 406 E-----TPYNIDCSHQ-----KSIL-----TOIPGAR----- 427
Db 567 KNGLTPLHVAHVHNNLDIVKLLPRGGSHPSPAWNGYTPPLHIAAKQNVARSLLQYGG 626
QY 428 -----HLSPTEDGMLGYDLYSSALADILSEPTWQ-----PIC-- 462
Db 627 SANAESVQGVTPHLAAQEGHAEMVALLSKQANGLNKSGLTPLHVAQEGHVPADV 696
QY 463 -----VGLYAQWGSK--SFLKKLED-EMKTFAGQOQTEPLQFQSWL 501
Db 687 LIKHGVMDATTMRGVTPLHVASHYGNIKLVKELLQHQADVNAKTKLGY--SPLHQAAQ 744
QY 502 --IVFLTLLCGGLGVAPVD-----TNIAIATSLFLALIYFIYVYGGRRGE 553
Db 745 GHTDITVLLKNG-----ASPNEVSDGTTPLAIKRLGYISVTDVLKV----- 789
QY 554 SNWMAWLSRLARHIGYLELLKLFAMVNPPELPEQTKALPVRFLETDYNRLS-SVGG 612
Db 790 TDTSFVLSDKIR-MSFPTVDEILDVSEDEGEELIS-----FKAERRSRDVDEE 840
QY 613 TSAEMIATLSDACEREFGLATRLPR-----VFRTEE-SQKKKWKTKCCLPSFV 662
Db 841 KELLDFVPKLDQVVE-----SPAIPRIPCAMPETVIRSEEQEASKEYDEDSLIPS-- 892
QY 663 IFLFIVGCIITAGITLAI FRVDPKHLTVNAILLSIASVUGLAVLNCRTWWQVLDLNS 722
Db 893 -----SPATSDNI-SPVASPVHTGFLVSF-----MVDARGGS 925
QY 723 QRKRLHSAASKHLKLSGKPMVKCEVELMARMKATIDFTQNTQRLVLIIDGLDACEQ 782
Db 926 MRGRSHN-----GURVIP-----PRCAAPTITRLV----- 954
QY 783 DKVLQMLDTRVRLFSKGPPIAFASDPHIIIKAINQNLSVLNDRSDNNGHYMRN-IVHL 841
Db 955 --KPQKLSLTPPL-----ABEEGLASRIITAGTGAQFLSPVIVEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSQNSLGEWTKLGSKTALN 900
Db 994 PHFASHRG---DRELVLVLRSENGV-----WKEHRSRYGESYLDQILN-GMDEELG 1041
QY 901 RRDYTRRQOMORTITROMSFDLTKLVT--EDWFSDISPQTMRLNINIVSVTGRLLRAN 957
Db 1042 SLELEKRVCRIT--TDPLFVMSRLCQD-YDIIGEG-----GSLSKVLPLV 1091
QY 958 QITFNMDRLASWINLTQWPYRTSWLILYLEETEGLPDQMTLK-----TWYER 1005
```

```
Db 1092 QATFPENAVTKRVKLAQ-----AQVPVDELVTKLQGNQATFSPIVTVPR 1137
QY 1006 ISK-----NIPTKDVEPLEIDGDIRFEV-----LSRTPVLVA 1042
Db 1138 RRKFHRPIGLRIPLPPSWTNDPRDSGEGDTLSRLLCVSVIGGTDQAWEDITTKLVYA 1197
QY 1043 RDVKTEPLCTVNLDPKRLRIIADVRAAREQINIGLAYPPLPLHEGPPRPPSGYSPASV 1102
Db 1198 NECANF--TTNV--SARFWLSDCPTAEAVNFATILLYKEL-----TAVPYMAKF 1242
QY 1103 CSASFNPGPPGWSVPQPHSSYYSGLSGFQHPFYNRAAVPATGSSLLSSMTVDVCEK 1162
Db 1243 VIFAKMNDPREG-----RRCYCMTDDKVDKT 1269
QY 1163 LRQ-----LEG-----LDQNMMPQYCTTIKAN-----INGRVL 1191
Db 1270 LEQHENFVARSRDIEVLEGMSLFAELSGNLVP-----VKAAQQRSHFQSFRENRLA 1324
QY 1192 SQCNIDELKKE-----MAMNFGD-WHLFRSMVLEMRSVESQVVPEDPR----- 1233
Db 1325 MPVKVRDSDREPGGSLSLFLRKAMKYEDTOHILCHLNIWPPCAKSGAEDRRRTPTPLAL 1384
QY 1234 ---FLNENSAPVPHGESARR-----SSHELPLTELSSQTPYTLNFSFEELNTGLD-- 1283
Db 1385 RYSILSESTPGSLSGTEQAEMKMAVISEHLGLSWAELARE---LQFSVEDINRIRENVP 1440
QY 1284 ----EGAPHSNLSWOSQTRTPSLSSINSQSSISLTKDKVQA-----EYRD 1329
Db 1441 NSLLEQSVALLNL-WIREGONANMENLYALQSIDRGEIVNNLEGSGRQSNLKPDRRH 1499
QY 1330 AVREYIAQMSQLEGGTSGSTISGRSPHSTYVYGSSSGSIHSTLEQREGKELQED 1389
Db 1500 TUDYLSLSPQWNG--YSSLQDELSP-----ASLGCALSSPLRADQ----- 1539
QY 1390 GRKSFMLKRGDVIDYSSSGSVNNEASPLDPI-----TEED---EKSQOS--GSKLLPKCKS 1440
Db 1540 -----YMWAVLDAIPLAATEHDTMLEMSDMQVMSAGLTPSLVT 1579
QY 1441 SERPSLFTQDLKLGGLRYQKLPSEDESGTGRVQITPHCSKWIRTKLKAQRECAP 1500
Db 1580 AEDSSL-----ECSKAEDSDATGH-----EWKLEGALESEPRGP 1613
QY 1501 QEHSABPIRTFIKAEYLSDA---LLDKKSDSDSGVRSNESSPNHSHNEAADDQLEKA 1557
Db 1614 ELGSLR---LVEDDVTDSDATNGLIDLLE-QEKGORSEKLPFGSKRQDDATGAGQ-DSE 1667
QY 1558 NLIEDEDEHSGKRGMPHSLGSLQDPIIARMSICSEDK-----KSPSECSLIASSPEESW 1612
Db 1668 NEVSLVSGHQRGQARITHS-----PTVSQVTERSQDRLQDWDADGSIYSYLQDAAGSW 1721
```

RESULT 5

AE2149

hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)

A;Species: Nostoc sp. PCC 7120

C;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AE2149

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yaeuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab

A;Reference number: AB1807; UID:21595285; PMID:11759840

A;Accession: AE2149

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-426 <KUR>

A;Cross-references: UNIPROT:Q8YTG9; GB:BA000019; PIDN:BA74447.1; PID:g17131841; GSPDB:G

A;Experimental source: strain PCC 7120

A;Genetics:

A;Gene: all2748

Query Match 6.3%; Score 555; DB 2; Length 426;

F;601-633/Domain: ankryrin repeat homology <AN18>	
F;634-666/Domain: ankryrin repeat homology <AN19>	
F;667-699/Domain: ankryrin repeat homology <AN20>	
F;700-732/Domain: ankryrin repeat homology <AN21>	
F;733-765/Domain: ankryrin repeat homology <AN22>	
F;766-798/Domain: ankryrin repeat homology <AN23>	

Query Match 6.2%; Score 553; DB 2; Length 1856;
 Best Local Similarity 20.5%; Pred. No. 8.8e-24;
 Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;

Qy	21	LKALLECKDVRBERNCCQTPLMAAEQGNVEIVKELLKNGANCUEDLDNWTALISASK	80
Db	94	VRELNVYGANNVAQSOKGFTPLYMAAQENHLEVVVFKLLENGANGVNATEDGFPTPLAVALQ	153
Qy	81	EGH-----	83
Db	154	OQHENVVAHLINVTGGKVRLPALHIAARNDDTRTAAVLLQNDPNPDVLSTGTPTLHIA	213
Qy	84	----IHIVELLKSGASLEHRDMGWGWTALMWAACYKGRTDVWELLSHGANSPTSGLQYS	138
Db	214	AHYENLVAAQLLNRGSSVNFTPONGITPLHIASRREGNVIMVRLLLDRGAIE-TTKYDE	272
Qy	139	VYPFIWAAGGHADI VHLLLONGAKVNCSDKYTTTPLVWAARKGHLECVKHLLAMGADVD	198
Db	273	LTPHCAARNGHVRISEILLDHCAPIQAOKTKNGLSPHMAAQGDHLDCLVRLLLQYDAEID	332
Qy	199	-----OEGA-----NSMTALI VAKGYGTSQSVKEILK	225
Db	333	DI TDLHLTPLHVAACHGHHRVAKVLDKGAKPSNRALNGFTPLHIACKNHVMRMELLK	392
Qy	226	RNPVNVLTDKGN TALMIASKEGHEIIVODLLDAGTYVNI PDRSGDTVLIGAVRGGHVEI	285
Db	393	TGASIDAVTESGLTPLHVASFPGHLPFIVKNLLQRGASPVNSVKETPLHMAARAGHTEV	452
Qy	286	VRALLQYADIDIRGDNK TALYWAYEKGNATMVWRDILQNCPDTEICTDKGETPLIKATK	345
Db	453	AKYLLQNKAKNAKADQDTPLHCAARI GHTNMVKLLENNANPNLAT TAGHTPLHIAAR	512
Qy	346	MRIEIVEVELLLDGAKVASVDKKGDTPLHVAINGRSRRLAELLRNPKDGRLLYRPNKAG	405
Db	513	EGHVETV LALKEASQACWCTKKGFTPLHVAKYGVKRVAEELLE---RDA---HPNAAG	566
Qy	406	E---TPYNIDCSHQ-----KSII-----TQIFGAR-----	427
Db	567	KNGLTPLHVAVHHNL D I VKLLP RGGSPHS PANWG Y TPLHIAKONQNEVARSLLOYGG	636
Qy	428	-----HLSPTETDGM LGDYLYSSALADIL SEPTMQP-----PIC--	462
Db	627	SANAESVOGYT PLHLAAQGEHAEWALLSKQANGNLGNKSGITPLHLVAAQE GHVPVADV	686
Qy	463	-----VGLV AQWSGK - SP LLKKLED - EMKT FAGOOTEPLPFQFSWL	501
Db	687	LKHGWVMDATT RMGYT PLHVASHYGN IKLVKFP LLQHQADVNAKTKLG -- SPLHQAAQQ	744
Qy	502	--IVFLTLLCGGLGVFAFPVD-----TNLAIALSFLAIYIFFVI VIFGGRGE	553
Db	745	GHTD IVTL LKNG-----ASP NEVS DGDTPLAI AKRLGIYSVTDVLKW-----	789
Qy	554	SNWNWALSRLARHIGYLELLFKMLFMVNPPELPEQT KALPVRFLUFDYNNRLS-SVGGE	612
Db	790	TUETSPLVSDKHR-MSFPETVDEILDVSEDEGEELIS-----FKAE RRDSRDVDEE	840
Qy	613	TS LAEMIATLSDACREDFGLATRLFR-----VFRT EE -SQGKKWKTKCCLPSFV	662
Db	841	KELLD FPVKLDQWE-----SPAIPRICAMPETVVISEREQBASKEYDEBDSLIPS--	892
Qy	663	I FLFVGCIIAGITL LAIFRVPDPKHLTVNALISTASVGLAFVLNCRTWQVLD SLINS	722
Db	893	-----SPAT ESDNI -SPVASPVHTGFLVSF-----MVDARGGS	925
Qy	723	QRKRLHSAASKLHKLS EGFMVKLCLEVELMARMAKTIDSFTQNQTRLVVIIDGLDACEQ	782

Db 926 MRGRHN-----GLRVIP-----PRTCAAPRTICELV----- 954
Qy 783 DKVQLMDTVRVLFSGKPFIAFASDPHIIKAINQNLNSVLDRSDNNGHNDYMRN-IVHL 841
Db 955 --KPQKLSPPPL-----AEEGLASRIITAGTGAQFLSPVIVEI 993
Qy 842 PVFL-NSRGLSNARKFLVTSATNGDITCSITGTQEDTDRVSQNSIGENTKLGSKTALN 900
Db 994 PFAHSGF--DRELVLVRSNGSV-----WKEHRSRYGESYLDQILN-GMDELG 1041
Qy 901 RRDYRRROMORITROMSDLTAKLVLT--EDWFSDISPQTMRLNLINIVSVTGRLLRAN 957
Db 1042 SLELEKRVCRIT--TDPLFVMSRLCQD-YDIIGPEG-----CSLAKSLVPLV 1091
Qy 958 QITFNWDRLASWNLTEQWYRTSWLILYLEETEGLPDQMTLK-----TWYER 1005
Db 1092 QATFPENAVTKRVKALQ-----AQVPDELVTLLGNQATFSPIVTVEPR 1137
Qy 1006 ISK-----NIPTKDVEPLEIDGDIRNEVP-----LSRTPVLVA 1042
Db 1138 RRFPHRPIGIRIPLFPSTWNTNDRSDGEGDTLSRLLCVIGGTDQAQWEDITGTLVYA 1197
Qy 1043 RDVKTFPCTVNLDPKLEIADVRAAREQINIGLAYPPLPLHEGPPRPSPGVSQPASV 1102
Db 1198 NECANF--TNNV--SARFWSLSDCPTAEAVNATLLYKEL-----TAVPYMAKF 1242
Qy 1103 CSASAFNGPPFGVSPQSHSYVGLSGPQHPFYRNAAPVATGSSLLLSMTVDVCEK 1162
Db 1243 VIFAKMDPREG-----RLRCYCMTDDKVDKT 1269
Qy 1163 LRQIEGLDQNMPOYCTTIKK-----ANINGRVLSCQNIDELKEMANFGDHLFRS--M 1216
Db 1270 LEQHENEVEVARSDIEVLEGLMSLFAELSG-----NLVPVKAQAQORSFPHOSFRENRL 1323
Qy 1217 VLEMRSVESQVP-----EDPRFLNENSAVP-----HGESARSSHTLP 1258
Db 1324 AMPVKVDSRREGGSLSPFKAMKYEDTQHILCHLNTWPPCAKSGAEDRRTPPTPLA 1383
Qy 1259 L--TELSSQPTYTINFSFE-ELNLTGLDEGAPHNSLNWSQSTRTP-SLSSLSNQDSII 1314
Db 1384 LRYSLSESTPGSLSGTEQAEKMAVISE-----HLGLSWAELARELOQFVEDIN----- 1433
Qy 1315 EISKLTQVQAEYDAREYIAQMSQLEGGTSGSTISGRSSPHSTYVI--GQSSSGSGSIH 1372
Db 1434 -----RIRVENPNSLLEQSVALLNL-----WVIREGQANMENLY 1468
Qy 1373 STLEQ-ERKEGELQBDGRKSLFKM--RGVDIYSSGVSVTNEASPLDPITEEDR-- 1425
Db 1469 TALQSIDRGEIVNMLESGRQSRNLKPDRRHTRDYSLSPSQMGYSYL-----QDELLS 1523
Qy 1426 -----KSDQSGSKL-----LPGKKSERPSLFQTDLLKLGGLRYQKLPDDED 1468
Db 1524 PASIGCALSSPLRADQYWEVAILDAIPLAATEHDTWLEMSDMQVWSAGLT-PSLVTAED 1582
Qy 1469 ESGTGRVQIIPHCSSMIRTKLAKQECASQBSAEPTRTFKAXEYLSDALDKDS 1528
Db 1583 SS-----LECSK-----AESDATHGEWKLEGALSEEPGPELSLELVEDDTVDSDAT 1631
Qy 1529 S-----DSGVRASNESSPNHSLHNEAADDQLEKANLIELEDEHSGRGRMPHSLSLG 1580
Db 1632 NGLDILLEQEBGQSRSEKLPQSKRQDDATGAGQ-DSNEVSLVSGHQRGQARITHS----- 1686
Qy 1581 QDPIIARWICSEDK-----KSPSECSLIASSPESW 1612
Db 1687 --PTVSQVTSQRDLQWDADGSIYSYLODAAQGSW 1721

RESULT 7

A35049

N; erythrocyte splice form 2 - human

N; Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R

N; Contains: ankyrin 2.2, erythrocyte

C; Species: Homo sapiens (man)

C; Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 09-Jul-2004
C; Accession: A35049
R; Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; Ke
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A; Title: cDNA sequence for human erythrocyte ankyrin.
A; Reference number: A35049; MUID:90175370; PMID:1689849
A; Accession: A35049
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-1880 <LAM>
A; Cross-references: UNIPROT:P16157; GB:M28880
C; Genetics:
A; Gene: GDB:ANK1; ANK
A; Cross-references: GDB:118737; OMIM:182900
A; Map position: 8p11.2-8p11.2
C; Superfamily: ankyrin; ankyrin repeat homology
C; Keywords: alternative splicing; cytoskeleton
F; 2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F; 2-1513/1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F; 44-76/Domain: ankyrin repeat homology <AN01>
F; 77-109/Domain: ankyrin repeat homology <AN02>
F; 110-142/Domain: ankyrin repeat homology <AN03>
F; 143-171/Domain: ankyrin repeat homology <AN04>
F; 172-204/Domain: ankyrin repeat homology <AN05>
F; 205-237/Domain: ankyrin repeat homology <AN06>
F; 238-270/Domain: ankyrin repeat homology <AN07>
F; 271-303/Domain: ankyrin repeat homology <AN08>
F; 304-336/Domain: ankyrin repeat homology <AN09>
F; 337-369/Domain: ankyrin repeat homology <AN10>
F; 370-402/Domain: ankyrin repeat homology <AN11>
F; 403-435/Domain: ankyrin repeat homology <AN12>
F; 436-468/Domain: ankyrin repeat homology <AN13>
F; 469-501/Domain: ankyrin repeat homology <AN14>
F; 502-534/Domain: ankyrin repeat homology <AN15>
F; 535-567/Domain: ankyrin repeat homology <AN16>
F; 568-600/Domain: ankyrin repeat homology <AN17>
F; 601-633/Domain: ankyrin repeat homology <AN18>
F; 634-666/Domain: ankyrin repeat homology <AN19>
F; 667-699/Domain: ankyrin repeat homology <AN20>
F; 700-732/Domain: ankyrin repeat homology <AN21>
F; 733-765/Domain: ankyrin repeat homology <AN22>
F; 766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 553; DB 2; Length 1880;

Best Local Similarity 20.5%; Pred. No. 9e-24; Mismatches 648; Indels 574; Gaps 73;

Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;

Qy 21 LKALLEKCDVDRNECGQTPLMAAEQGNVEIVKELLKNGKNCNLEDDLNTALISASK 80
Db 94 VRELNVYGANVNAQSQKGFPLYMAAQENHLEVVLKNGANQNVATEDGFTPLVALQ 153
Qy 81 EGH----- 83
Db 154 QGHENVVAHLINYGTKGVRLPALHIAARNDDTRTAALLQNDPNPDVLSKTGFTPLHIA 213
Qy 84 -----THIVEELKSGASLEHRDMGWTALMAWACYGRDVTVELLSHGANSPTGLQYS 138
Db 214 AHYENLVQAQLLRSSVNFPTQNGITPLHIAARRGNVIVWRLLLDGRQAIE-TTKYDE 272
Qy 139 VYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPPLVWAARKHLECVKHLAMGADVD 198
Db 273 LTPHCAARNGHVRISSEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCCVLLQYDAID 332
Qy 199 -----QEGA-----NSMTALIVAKGYTQSVKEILK 225
Db 333 DITLDHLTPLVAAHCHGHRVAKVLLDKGAKPNRSALNGFTPLHIAACKXNHRVMEILLK 392
Qy 226 RNPVNLTQDKGNTALMIASKEGHEIVODLLDAGTVVNIPIPRSDGTVLIVAGRGHVEI 285
Db 393 TGASIDAVTESGLTPLHVASFMGHLPIVKNLLOGRASPNSVNVKVFPLHWAARAGHTEV 452
Qy 286 VRALLQKYADIDIRGQDNKTALYWAVEKGNATVWRDILQCNPPTEICTKDGFTPLIKATK 345

Db 453 AKYLONKAKNAKADQDTPLHCAARIHTMTNMVLLLENNANPLATTAGHTPLHIAAR 512
QY 346 MENIEVELLDKGAUVKADKGGDTPLHVAIRGRSRLAEALLRNPKDGRLLYRPNKAG 405
Db 513 EGHVETVLALLEKASQACWKTKGFTPLHVAKYKVRVAELLLE-RDA--HPNAAQ 566
QY 406 E--TPYNIDCSHQ-----KSTL-----TQIFGAR----- 427
Db 567 KNGLTPLHVAVHNHNDLVKLLPRGSPHSPANNGYTPHLIAAKONQVEVARSLQYGG 626
QY 428 -----HLSPTETDGMGLDYLYSALADIILSEPTMQP-----PIC-- 462
Db 627 SANAESVOGVTPLHAAQEGHAEMVALLLSQANGNLGNKSGLTPLHLVAQEGHVPVADV 686
QY 463 -----VGLYAQWGSCK--SFLIKKLED-EMKTFAGQTEPLFQPSWL 501
Db 687 LIKHGVMDATRMGYTPPLHVAHYGNIKLVKFLLOHQADVNAKLYG--SPLHQAQQ 744
QY 502 --IVPLTLLCGGLGVAFPVD-----TNLAIAISLFLAIYIPFIVYFGRREGE 553
Db 745 GHTDVTLLKNG-----ASPNEVSSDGTTPLAIAKRLGYISVTDVLKV-- 789
QY 554 SNNAWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRFDTYRNL-SVUGE 612
Db 790 TDTSFVLVSDKHR--MSFPETVDEILDVSEDEGEELIS-----PKAERDRSDVDEE 840
QY 613 TSLAEMIATLSDACERBERGFLATLFR-----VPTTEE-SQGKKKKKCTCLPSFV 662
Db 841 KELLDVFKLQOVV-----SPAIPRCPAMPETVIRSEEQASKEYDEDSLIIPS-- 892
QY 663 IFLFVIGCIIAGITLLAIFRVDPKHLTVNAILIISIASVVGFLAVFLNCRNTMQVLDLSLNS 722
Db 893 -----SPATETSDNI-SPVASPVHTGFLVSF-----MVDARGGS 925
QY 723 QRKRLHSAASKLHLKSGFEMVKVKEVELMARWAKTIDSTQNTQRLVLIIDGLDACEQ 782
Db 926 MRGSRHN-----GLRWIP-----PRTCAAPTRITCRVL----- 954
QY 783 DKVLQMLDTRVLVFSKGFPIAFASDPHIIKAINQNLNSVLDRNSINGHYMEN-IVHL 841
Db 955 --KQKLSSTPPL-----ABEGLASRIIALGPTGAOFLSPVIVEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITGSDTTGTQEDTDRRVSONSLGEMTKASKTALN 900
Db 994 PHFASHGRG---DRELVLVRENSGV-----WKEHRSYGESYLDQILN-GMDEELG 1041
QY 901 RRDYRRQMQRTTROMSFDLTKLVLT---EDWFSBISPTQMRRLNIVSVTGRLLRAN 957
Db 1042 SLELEKKRVCRITIT--TDFFLYFVIMSRQCQD-YDIIGPRG-----GSLKSKLVPLV 1091
QY 958 QITFNWDELASWINLTQWPVRTSWLIIYLEETEGLPDQMTLK-----TWYER 1005
Db 1092 QATFPENAVTKRVKALQ-----AQVFPDELVTKLGNQATFSPITVTEPR 1137
QY 1006 ISK-----NIPPTKQVPELLEIDGIRNPEVF-----LSSRTPVLVA 1042
Db 1138 RRKHFRPGLRILPBPSTWNPDRSGEDTSLRLCSVIGGTQAOQWEDITGTTKLUYA 1197
QY 1043 RDVKTFLPCTVNLDPKLEIIRIADVRAAREQINIGLAYPPLHEGPPRPPSGYSOPASV 1102
Db 1198 NECANF---TTNV--SARFMLSDCPRTAEAVNFATLLYKEL-----TAVPYMAKF 1242
QY 1103 CSSASFNGPPPGVGVSPQPHSYYSGLSGPOHPFYRNAAPATGSSLLLSMTVDVCEK 1162
Db 1243 VIFAMNDPREG-----RLRCYCMTDKDYKT 1269
QY 1163 LRQIEGLDQNMPOYCTTIKK-----ANINGRVLSQCNIDELKEMAMNFGDHLFRS--M 1216
Db 1270 LEQHENFVEVARSDIEVLEGMSLFAELSG-----NLVPVKAAQORSFHQSFRENRL 1323
QY 1217 VLEMRSVESQVPP-----EDPRFLNENSSAPVP-----HGESARRSSHTPLP 1258
Db 1324 AMPVKVRDSSREPGLSFLRKAMKYEDTQHILCHLNIITWPPCAKGSAGDEDRRTTPLA 1383

QY 1259 L--TELSSQTPVTLNFSEF-ELNTLGLDEGAPRHSNLSWQSQOTRTP--SLSSLSQDSSI 1314
Db 1384 LRYILSESTPGSLSGTEQAEMKAVISE-----HLGLSWABLARELQFSEVDIN----- 1433
QY 1315 EISKLTQKVAERYDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYI--GQSSSGGSIH 1372
Db 1434 -----RIRVENPNSLLEQSVALLNL-----WVIREQONANMENLY 1468
QY 1373 STLBO-ERKKEGELKQSDGRKSFMLK---RDVIDYSSSGVSTNEASPLDPITEDE--- 1425
Db 1469 TALQSIDRGEIVNMLEGSGRNLKPDRRHTDRYDLSLPSQMGYSYL-----QDELLS 1523
QY 1426 -----KSDQSGSKL-----LPGKKSSRPSIFQTDLKLKGGGLAYQKLPSDED 1468
Db 1524 PASLGCALSSPLRADQTNVEVAIILDAIPLAATEHDTMLEMSDMQVMSAGLT-PSLVTAED 1582
QY 1469 ESGTGRVQITPHCSKMIRTKLKAQREKASPOHSAEPIRTFIKAKELYLSDALLDKKDS 1528
Db 1583 SS-----LECSK---AEDSDATGHEWKLEGALSEEPGPELGSLELVDDTVDSDAT 1631
QY 1529 S-----DSGVRSNESSPNHSLHNEAADSQLEKANLIELEDEHSGHGRKGMPHSLSGL 1580
Db 1632 NGLIDLLEQEBGQRSEKLPQSKRQDDATGAGQ--DSENEVSLVSGHQRGQARITHS----- 1686
QY 1581 QDPITIRMSICSDEK-----KSPSECSLIASSPEESW 1612
Db 1687 --PTVSQVTERSQDRLQDMDADGSIVSYLQDAAGGSW 1721

RESULT 8
T42716
ankyrin 3, splice form 4 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42716
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: Z22237; MUID:95340633; PMID:7615634
A:Accession: T42716
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1961 <P>
A:Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 6.2%; Score 552; DB 2; Length 1961;
Best Local Similarity 20.2%; Pred. No. 1.1e-23;
Matches 427; Conservative 289; Mismatches 674; Indels 724; Gaps 85;

QY 14 EENIPALKALKKCKVDVERNECGQTPLMAAEEQGNVEIVKELLKNGANCNLEDLDNWT 73
Db 66 KEGHVEVSELLQREANVDAATKKNTALHIASLAGQAEVVKLVITNGANVNAQSQNGFT 125
QY 74 ALISASKEGHIHVEELLSKASLEHRDMCGWTALMWACYKGRDVTVELLL----- 124
Db 126 PLYMAAQENHLEVVRFLLONGASQSLATEDGFTPLAVALQQHQDVVSLLENDTKGVR 185
QY 125 -----SHGANPSVTGL----- 135
Db 186 LPAHLTAARKDDTKAAALLQNDTNADVSKSGFTPLHIAAHYCNINATILLNRAAAYD 245
QY 136 ---QYSYVPIIWAAGRHADIVHLLLQNGAKVNCSDKYGTTPLVWAARKG----- 182
Db 246 FTARNIDITPLHVASKRGNANWVKLLDRGAKIDAKTRDGLTPLHCGARSQHEQVQVEMLLD 305

183 QY -----HLECVKHLAWGADVDOEGANSMTALIVAKGYTQS 219
306 Db RSAPILSKTKNGLSPLHMATQGDHLCVQLLOHNPVDDVTNDYLTALHVAACHGYKV 365
220 QY -----VKEILKRNPNVNTDKDGNALTAMIASKE----- 247
366 Db AKVLLDKASPNKALNGFTPLHIAKKNRIRVWELLKKGASIQAVTESGLTPIHVAAF 425
248 QY -GHIEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRQNKTA 306
426 Db MGHVNIQSOLMHHGASPTNTNRGETALHMAARSGQAEVRYLVQDQAQVEAKAKDQTP 485
307 QY LYNAVEKGNATMRDILLOCPDTEICTKQGETPLIKATKRNIEVVELLLDKAKYSADV 366
486 Db LHSARLGKADIYQQLQOGASPNAAATTSYTPLHLAAREGHEDVAFLDHDGASUITT 545
367 QY KKGDTPLHVAIRGSRRLAELLAR-----NPKDGRLL----- 398
546 Db KKGFTPLHVAKYKLEVASLLQKSASPDAGKSGLTPLHVAHYDNQKVALLLDQGA 605
399 QY --YRPNKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPTETDGMGLGY 441
606 Db SPHAAKNGYTPLHIAKKNQMDIATSLLEYGADANAVTRQGIASVHLAAQEGHVDMSVL 665
442 QY DLYSSALADILSEPTWOP-----PICVGLYAQM 469
666 Db LLSRANVNLSNKSGLTPLHLAQEDRVNVAEVLVNOGAHVDAQTKMGYTPLHVGHCH--Y 723
470 QY GSGK--SFLIK--KLEDEMKTFAGQOQTEPLFQSMILVFLTL-----LLCGGLGLV 516
724 Db GNKIVNFFLOHAKVNAKNGVGTALHQAQQGHTHINVLLQNNASPNELTVNG--- 779
517 QY FAPPVDTNLAIASLSFLALIIYFFIVI----- 544
780 Db -----NTALAIARRLGIYVVDTLKVTEBIMTTTTITEKHKMNPETWNEVLDMSDEV 834
545 QY -----YFGRREGESNNAWALSTRLARHIGYLELLFKL-----MFVN 582
835 Db RKASPEKLSGGEYISDGBGDKCTWPKIPKVQ-----EVLKSEDAITGDKILG 886
583 QY PPELPEQTTKALPVRFLEFDYNNRLSSVGGETSALMIATLSDACERFEGFLATRLFRVFR 642
887 Db PDLKELGDDSLPAE-----GYVG--FSLGARSASLR-----SFSSDRSYTLNR 928
643 QY TBSQGGKKKKWKCCLPSFVIFLPIVGCIIAGITLTLAIFRVPD-KHUT-----VNAI 693
929 Db SSVAR-----DSMMIEELLVPSKEQHLTFTREFDSDSLRHYSWAADTLDNVNLV 977
694 QY LISTASVVLGAFVLNCRWQVLDLSLNSQ-----RK-----RLHSAASKLHKL- 737
978 Db SSPVHSGFLVSFWMVDARG-----GSMGRSGRHGMRRIIPPRKCTAPTRITCRLVVRHKL 1032
738 QY -----KSEG-----FMKVLKCEVELMARMAKTIDSFON 766
1033 Db NPPMVBEGGLASRLVEMGAPAGQFLGPVIVEIPHFGSMRGERELIVLRSENGETWKEH 1092
767 QY -----TRLVVIDGLDACEQDKVQLMDTVRV--LFSKGPFTAIAPASDHIHIIKAINQN 819
1093 Db QFDSKNEDLAEELNGMD--BELDPSPEELGTKRICRIITK-DFOQYFA-----VVSRIKOE 1144
820 QY LNSVLRSNNGHDYMENIVHLPVFLNSRGLSNARKFLV-TSANGDITCSDTTGTQ--- 875
1145 Db SNQIGPSEGGI-----LSSTTVFLVQASFPPEGALTKRIRVGLQAOP 1184
876 QY --EDTDRRVSONSLGEMTKLGSKTALNRDITY--RRRQMORTITROMSFOLTKLL-----VT 928
1185 Db VPEETVKKI-----LGNKATFSPIVTEPRRRKFKHPIT--MTIPVPPSPGEGVS 1232
929 QY EDWFSDSIPQMRLLNIVSVTG-----RLLRANQITFNWDRLASWINLTQWYPT 980
1233 Db NGYKGDATPNL--RLL--CSITGTSQAQWEDITGTTPLTFIKDCVGFNTNVSARF--- 1284
981 QY SWL-----ILYLETEGLPDQWTLKTM-YERISKNIPTTKDVEPLLEIDGDIRNF----- 1029

1285 Db -----WLAACHQVL---ETVGLASQLYRELI CVPYNAKFVFAKTWDP-----VSSLRCFCQWTD 1337
1030 QY -----EVLSSRTPLVARVDKTELPCTVNLDPKLR-----EIIADVARARE-- 1071
1338 Db RVDKLTLEQOENFEVARSKDIEVLEKPI--YVDCYGNLAPLTGKGQQLVFNFSYSPKRN 1395
1072 QY ---QINIGGLAYPPLP---LHEGPPRPFGSGYSPASVCSSASFNGFPFGVVPSPQHSS 1124
1396 Db LPSIKIRDTSOEPCGRLSFLKE--PKTTKGLUPOTA--VCN----- 1432
1125 QY YSGLSGPQHPFNRAAVPATGSSILL-----SMTVDVVCEK--LRQIEGLDQNM 1173
1433 Db --LNIITLPAHKAEXADRQSFASLALRKYSVLTEPSMSPQSCERTDIRMAIVAD-HL 1489
1174 QY MPOYCTTIKANINGRVLSQCNIDELKEMAMNF--GDW-----HLFRSVMLE 1220
1490 Db GLSWTELARELNFSDVINQIRVENPNLSISOSFMLLKWVTRDGNATTDALTSLVTKI 1549
1221 QY RSVESQVPEDP-----RFLNENS--SAPVPHGESARRSSHTLPLTSLSSOTP 1267
1550 Db NRIDIVTLLEGPIFDYGNISGTRSFADENNVDHV-----DGH--PSFOVELETP 1598
1268 QY YTLNPFSEELNTLGLDEGAPRHSNLSWQSQTRTTPSLSS-----LN 1308
1599 Db MGLYWT--PPNPPQDD--HFSDISSIESPFTPSRLSDGLVPSQGNIEHPTGPPVVT 1653
1309 QY SQDSSIEIKLTDKV---QAEYRDAYREIVIAQMSOLEGTSSTISG-----RSP-- 1356
1654 Db AEDTSLDSKMDSDSVTVTPADPLDVDESQKDLQSECAQCWASVPGIPNDGRQAEPLR 1713
1357 QY HSYIYIGQSSGSIHSTLEQERKQEGELKQEDGRKSFMLKRGDVIDYSSSGVSTNEASP 1416
1714 Db PQRKVMGSE-----QOEKRGSG-----P 1733
1417 QY LDPITBEDSKDSQSGSKLLPGKKSRPSLFTQDLKLGGLRYQKLPSDESGTGRVQ 1476
1734 Db DEEVTEKVK-----SLFE-DIQLE--EVEAEEMTEQOGAMLNRVQ 1772
1477 QY ITPHCSKMRITKLLKAKQREKASQPHESAEPIRTFFIKAEYLSDDLKDKSDSGVRSN 1536
1773 Db -----RAELAMSSLAGWNETPSSLESQAARR-----LTGGLLDRLDSSDQARDS 1820
1537 QY -----BSSPNHSLHNEAADDQLEKANLIELEDEHSGKRGK-----PHSL 1577
1821 Db ITSYLTCGPKIBANGNHT-----AEVPEAKAKPYFPESQNDIGKQSIKENLKPETHGC 1875
1578 QY SGLQDPIIARMSICSDKKSPPSCS--LIASSPEESWPACQKAYNLNRPTSTVTLLNNTA 1635
1876 Db GRTEEPV---SPLTAYQKSLEETSCLVIEDAPKPCVPVGMKKM-----TRTT 1919
1636 QY PTNRANQNPDEIG 1649
1920 Db ADGKARLNLOEERG 1933

RESULT 9

T42715
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PRT>
A;Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:G710548; PID:G710549; PIDN:AB01604
A;Experimental source: strain C57BL/6J; kidney

Db 1541 EGPIDYGNISGTRSFADENNVPDPV-----DGH---PSQVELETWGLYWT--P 1587
Qy 1277 LNTLGLDGAAPHNLNSQSQTRTPSLSS-----LNSQSQSIEIS 1317
Db 1588 PNPFQDD--HFSDISIESPFRTPSLDGLVPSQGNIEHPGPGPPVTAEDTSLSDS 1644
Qy 1318 KLTKV-----QAEYRDAYREYIAQMSQLEGGTSGSTISS-----RSSP--HSTVYIQQS 1365
Db 1645 KMDDSVTVTDPADPLVDVDESQKDLQSECAQCAQWASVPGIPNDGRQAEPPLRPQTRKVGMS 1704
Qy 1366 SSGSIHSTLEQERKGEELQEDGRKFLMKRGVDVYSSGVSSTNDEASPLDPITEDE 1425
Db 1705 SE-----QOERKSGS-----PDEEVTEDEKV 1724
Qy 1426 KSDQSGSKLLPGKSSBRPSLFOTDLKLGGLRVOKLPSPDESGTGRVQITHCSCMI 1485
Db 1725 K-----SLFE-DIQLE--EVAEEMTEQGGQAMLNRVQ-----RAEL 1758
Qy 1486 RTKRLKAKQRECAQPOEHSABPIRTFIKAKELYLDALLDKKSDSGVRSN----- 1536
Db 1759 AMSSLAGQNETPSSGLESQAARR-----LTGGLDLRLDSSDQARDISITVLTGEP 1811
Qy 1537 ---ESSPNHSLHNEAADSQLEKANLIELEDEHSGKRGK-----PHSLSGLDPIIA 1586
Db 1812 GKIEANGHT-----AEVPEAKAKPPPEQSDNIGKQSIKENLKPXTHGCGRTEEPV-- 1864
Qy 1587 RMSICSEDKSPSCS--LIASSPEESWPAQCAKAYNLNRTSTVTLNNTAPTNRANQNF 1644
Db 1865 --SPLTAYOKSLEETSCLVIEDAPKPCVGVGMKKM-----TRTTADGKARLNL 1910
Qy 1645 DEIRG 1649
Db 1911 QEEEG 1915

RESULT 11
S37431
ankyrin 2, neuronal long splice form - human
N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 09-Jul-2004
C;Accession: S37431; B39643; A40334; A49462; S14569
R;Chan, W.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>
A;Cross-references: UNIPROT:Q01484; EMBL:Z26634; NID:G406287; PIDN:CAA81387.1; PID:G4062
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A;Reference number: A39643; MUID:91302466; PMID:1830053
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443,3585-3924 <OTT>
A;Cross-references: EMBL:X56958
R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
Genomics 10, 858-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92009921; PMID:1833308
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 463-474, 'PE', 477-495 <TSE>

A;Cross-references: GB:W37123; NID:G178647; PIDN:AAA62828.1; PID:G178648
R;Chan, W.; Kordeli, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A;Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and se
A;Reference number: A49462; MUID:94075409; PMID:8253844
A;Accession: A49462
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3924 <RES>
A;Cross-references: EMBL:Z26634; NID:G406287; PIDN:CAA81387.1; PID:G406288
C;Genetics:
A;Gene: GDB:ANK2
A;Cross-references: GDB:I27607; OMIM:106410
A;Map position: 4q25-4q27
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-3924/Product: ankyrin 2, long form #status predicted <NA>
F;2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA>
F;63-95/Domain: ankyrin repeat homology <AN01>
F;96-128/Domain: ankyrin repeat homology <AN02>
F;129-161/Domain: ankyrin repeat homology <AN03>
F;162-190/Domain: ankyrin repeat homology <AN04>
F;191-223/Domain: ankyrin repeat homology <AN05>
F;232-264/Domain: ankyrin repeat homology <AN06>
F;265-297/Domain: ankyrin repeat homology <AN07>
F;298-330/Domain: ankyrin repeat homology <AN08>
F;331-363/Domain: ankyrin repeat homology <AN09>
F;364-396/Domain: ankyrin repeat homology <AN10>
F;397-429/Domain: ankyrin repeat homology <AN11>
F;430-462/Domain: ankyrin repeat homology <AN12>
F;463-495/Domain: ankyrin repeat homology <AN13>
F;496-528/Domain: ankyrin repeat homology <AN14>
F;529-561/Domain: ankyrin repeat homology <AN15>
F;562-594/Domain: ankyrin repeat homology <AN16>
F;595-627/Domain: ankyrin repeat homology <AN17>
F;628-660/Domain: ankyrin repeat homology <AN18>
F;661-693/Domain: ankyrin repeat homology <AN19>
F;694-726/Domain: ankyrin repeat homology <AN20>
F;727-759/Domain: ankyrin repeat homology <AN21>
F;760-792/Domain: ankyrin repeat homology <AN22>
F;793-825/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 550; DB 2; Length 3924;
Best Local Similarity 20.4%; Pred. No. 4,78-23;
Matches 417; Conservative 279; Mismatches 727; Indels 626; Gaps 78;
Qy 14 EENIPALKALLEKCKVDVERNECGQTPMLAABQGNVEIVKLLKNGANCNLEDLNNWT 73
Db 73 KEGHVGVLVQELLGRLGGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAOQNGFT 132
Qy 74 ALISASKEGHIHIVEELLKSGAS-----LEHRDMG--- 103
Db 133 PLYMAAOENHIDVVVKYLLENGANGQSTATGFTPLAVALQGHNOQVAILLNKTGKVR 192
Qy 104 -----GWTALMWACYKGRDVVVELL 123
Db 193 LPALHIAARKDDTKSAALLQNDHNADVQSQMVNRTTESGFTPLHIAAHGVNVNVALTL 252
Qy 124 LSHGANPSVTGLQVSYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPPLVWAARKG- 182
Db 253 LNRGAADVFTA-RNGITPLHVASKRGNTNMVKLLLDRLGGQIDAKTRDGLTPLHCAARSGH 311
Qy 183 -----HLECVKHLAMGADVDQEGANSNTALIV 210
Db 312 DQVVELLERGAPLLARTKNGLSPLHMAAQGDHVECVKHLLOHKAPVDDVTLDTLTLV 371
Qy 211 AVKGYTQSVKEILKRNPNVNLTDKGNLTALMTASKS----- 247
Db 372 AAHCGHVTVTKLLDKRANFNARALNGFTPLHIAACKNRKIKVMELLVKYGASIQAITESG 431
Qy 248 -----GHIEIVODLLDAGTVYVNIPIRSQDVTILIGAVRGGHVEIVRALLQKYADID 297
Db 432 LTPIHVAAFMGHLNIVLLLLQNGASPDVTNIRGTALHMAARAGQVEVVRCLLRNGALVD 491

QY 1648 -EGIRETSQVILRPGSPNPNTA-----VQENLKSMHAKSQRSSYT 1688
Db 1751 QRGPETGSP--RAGKEPSLWAPESAFSQEVOGDELQNPGEQVTBEQFT 1798

RESULT 13
A55575
ankyrin 3, long splice form - human
N;Alternate names: ankyrin G
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 09-Jul-2004
C;Accession: A55575
R;Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A;Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax
A;Reference number: A55575; MUID:95138209; PMID:7836469
A;Accession: A55575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4377 <KOR>
A;Cross-references: UNIPROT:Q12955; GB:U13616; NID:G608024; PIDN:AAA64834.1; PID:G608025
C;Genetics:
A;Gene: GDB:ANK3
A;Cross-references: GDB:424503; OMIM:600465
A;Map position: 10q21-10q21
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: alternative splicing; peripheral membrane protein
F;73-105/Domain: ankyrin repeat homology <AN01>
F;106-138/Domain: ankyrin repeat homology <AN02>
F;139-171/Domain: ankyrin repeat homology <AN03>
F;172-200/Domain: ankyrin repeat homology <AN04>
F;201-233/Domain: ankyrin repeat homology <AN05>
F;234-266/Domain: ankyrin repeat homology <AN06>
F;267-299/Domain: ankyrin repeat homology <AN07>
F;300-332/Domain: ankyrin repeat homology <AN08>
F;333-365/Domain: ankyrin repeat homology <AN09>
F;366-398/Domain: ankyrin repeat homology <AN10>
F;399-431/Domain: ankyrin repeat homology <AN11>
F;432-464/Domain: ankyrin repeat homology <AN12>
F;465-497/Domain: ankyrin repeat homology <AN13>
F;498-530/Domain: ankyrin repeat homology <AN14>
F;531-563/Domain: ankyrin repeat homology <AN15>
F;564-596/Domain: ankyrin repeat homology <AN16>
F;597-629/Domain: ankyrin repeat homology <AN17>
F;630-662/Domain: ankyrin repeat homology <AN18>
F;663-695/Domain: ankyrin repeat homology <AN19>
F;696-728/Domain: ankyrin repeat homology <AN20>
F;729-761/Domain: ankyrin repeat homology <AN21>
F;762-794/Domain: ankyrin repeat homology <AN22>
F;795-827/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 526; DB 2; Length 4377;
Best Local Similarity 20.0%; Pred. No. 1.5e-21;
Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;

QY 14 EEENIPALKLECKOVDNERECQYPLMAAQGNVEIVKELKNGCANLELDLQNT 73
Db 83 KEGHEVVSSELLQREANVDAATKNGTALHTASLAGQAEVVKVLVTNGANVNAQSQNGFT 142
QY 74 ALISASKEGTHIVEELLKSGASLEHRDMGWTALMMAWACYKGRTDVVELL- 124
Db 143 PLYMAAQENHLEVVKFLDNGASQSLATEDGFTPLAVALQGHQVVSLLLENDTKGKVR 202
QY 125 -----SHGANPSVTGL----- 135
Db 203 LPALHIAARKDDTKAAALLQNDNNADVESKSGFTPLHIAAHYGNINVTALLNRAAAVD 262
QY 136 ---QVSVPTIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAARKG----- 182
Db 263 FTARNDDITPLHVASKRGANWVKLLDRGAKIDAKTRDGLTPLHCGARSQHEQVVELLD 322
QY 183 -----HLECVKHLAMGADVDEQANSMTALIVAVKGYTQS 219

Db 323 RAAPILSKTKNGLSPLHMATQGDHLNCVOLLQHNPVDDVTDNDYLTALHVAACHGYKV 382
QY 220 VTEILKRNPVNLTKDGNLTALMIASKE----- 247
Db 383 AKVLLDKKANPNNAKALNGFTPLHIAACKNRKIKMELLKHGASIOAVTESGUTPIHVAAP 442
QY 248 -GHIEIVQDLLDAGTVVNIPIRSDRGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTA 306
Db 443 MGVNIVSQMHGASPNNTNVRGETALHMAARSGAEVVRVLVDQGAQVEAKAKDDQIP 502
QY 307 LYWAVEKGNATVVRDILQ--CNPD-----EICT 333
Db 503 LHISARLKGADIVQQLQOGASPNATTSYGTPLHLSAREGHEDVAAFLDDHGASLITT 562
QY 334 KQGETPLIKATKRNIEVVELLDCAKYSADVKKGDTPLHVAIRGRSRLAEILLRNPK 393
Db 563 KKGFTPLHVAAYKGLVANLLQKSASPDAAKAGSLTPLHVAAHYDQNKVALLL---D 619
QY 394 DGRLLYRPNKAGETPVNIDC--SHQKSILTOI--FGAR-----HLSPTETDGM 438
Db 620 QGASPHAAAKNGYTPHIAAKKNQMDIATLLEYGADANAVTRQGIASVHLAAQEGHVM 679
QY 439 LGYDLYSSALADILSEPTMOP-----PICVGLY 466
Db 680 VSLLLGRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNGQAHVDAQTQKGYTPLEVGCH 739
QY 467 AQWGSCK--SEILK-KLEDEMTFAQOPELRFQFSLW--IVFLTLLCGGLGLVFAFP-- 520
Db 740 --YGNIKIVNFFLLQHSKAVNAKTNGY--TPLHQAQOQGHTHIINVLQNN-----ASPN 790
QY 521 ---VDTNLAIAISLSPLALIYIFFIVYFGRRGESNMAWALSTRLARHIGYLELLPK 577
Db 791 ELTVNGTALGI-----ARRLGYISVDT 814
QY 578 LMPV-----NPPE-----LPEQTTKALPVRLFTDYNRLSSV--GGET 613
Db 815 LKIVTEETMTTIVTEKHQNVPTWNEVLMSDDEVRKANAPEMLSGDEYISDVEGED 874
QY 614 SLA-----EMIAATLSDACERERFGLATL-----FRVFRTEESQKKKKWKTCC 658
Db 875 AMTGDDTKVLPQDLKELGDDSLPAGYMGFSLGARSASLSRFSFSSDRSYTLN--RSYAR 932
QY 659 PSFVIFLIVGCIAGITLLAIFRVDP--KHLT-----VNAILISIVSUGLAFVLC 709
Db 933 DSMIEELLVPSKEOHLTPTRFSDSLRHSWAADTLDNVNLVSSPIHSGFLVSEMVDA 992
QY 710 RTWQVLDLSLNSQRKLHSAASKLHKLXSEGFMKVL---KCEVELMARMAKTIDSTQ 765
Db 993 RG-----GSMRGSRHG-----MRIIIPRKCTAP-----TR 1019
QY 766 NQTRLVVIIDGLDACEQDKVLQMLDTRVRLFSKGPFIAPASDPHTIIKAINQNLNSVLR 825
Db 1020 ITCRLV-----KHKLAN-----PPH-----GERRGISRLV 1047
QY 826 DSNINGHDYMRN--IVHLPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSQ 884
Db 1048 EMGPAGAGLPGVIVPEIHPHSGMRG--KERELIVLSRSENGE-----TWKEHOFD---SK 1096
QY 885 NSLGEMTKL--GSKTALNRDVTYRRQMORTITRQMSFDLTLLVTEDEWPSDISPQTMRR 942
Db 1097 NE--DULELNGMDEELDSPEELGKRIICRIITKDF-----POYFAVVS--RIKQ 1142
QY 943 LLNIVSVTGRLLAN-----QITFNWDRLASWINLQEWFYRTSWLILYLEETEGLPDQM 997
Db 1143 ESNIQIPEGGILSSTTVPLVQASFPFGALTKRIRVGLQ-----AQPVDEI 1188
QY 998 TLKTM-----YERISKNP-----TTKQVEPLLEI 1022
Db 1189 VKKILGNKATFSPIVTVPEPRRRKPHKPIWTIVPPPSGEGVNGYKGDFTPNRLLC 1248
QY 1023 DG-----DIRNFEVFLSSRTPVLVARDVKTFPLCTVNLDPKLRBIIADVRAARQIN 1074

Db 1249 TGGTSPAOWEDI-----TGTTPLTFTKDCVSP-----TTNV--SARFWLADCHQVLETVG 1296
Qy 1075 IGLLAYPPLPLHEGPPRPGSGYSPASVCSASFNGPPGPGVVSQPQSHSSYYSGLSGPOH 1134
Db 1297 LATQLYREL-----ICV-----PYAKV-----VFAKNDP-- 1323
Qy 1135 PPYNRAAVATGSSLLLSMTVDVCEKLRQIEGLDQNMPOYCTTIKANINGRVLQOC 1194
Db 1324 -----VESSLRCFCMTDKDKTLEQOE----- 1346
Qy 1195 NIDELKEMAMFGDWHLFRSNVLEMRVESOVVPEDPFLN-ENSSAPVPHGE-----S 1248
Db 1347 NPEEV-----ARSKDIEVLEKPIYDVDCNTLAPLTKGGQOVLVN 1386
Qy 1249 ARRSHTTELP-----LTLESSOPPYTLNFSFELNLTGLDEGAPRNSLSWOSOTRTPSL 1304
Db 1387 FYSFKENRLPSIKIRDTISOEPCGRLSFLKERKTKGLPQAVCNLNTLPAHKETES- 1445
Qy 1305 SSLNSQDSISLTKDKVQAPYDAYRE---YIAQMSOLEGGTGSSTISGRSSPHSTYY 1361
Db 1446 -----DQDDEIE---KTRRQSPASLAKRYSYLTEPGMIERSTGAT---RSLP-TTV- 1492
Qy 1362 IQOSSGSGSIHLEQERKEGELQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDIT 1421
Db 1493 -----SYKPFSTRPYQSWTTAPITVPGPAKS-----GFTSLSSSSNTTPSAPLSIW 1541
Qy 1422 BEDEKSDQSGKLLPGKKSERPSLFTDLKLGGLRYQKLPSDEDESGTGRVOITPHC 1481
Db 1542 -----SVTSPISKSTL-----GASTTSSVKSISDV 1567
Qy 1482 SKMIRTKLAKQRECAPQEHASPIRTFKAEY-----LSDALLDK-----KDS 1528
Db 1568 ASPIRSLAT-----MSSPIKTVWSQSPYNIQVSSGTLARAPAVTEATPLKGL 1614
Qy 1529 SDSGVRSNESPNSHNEAADDLSQELKANLIELEDEHSGKRGPHS----- 1576
Db 1615 ASNTSFSTRSP-----VTTAGSLERSSTMTTPASPKNINMYSSLPFKSIITSA 1668
Qy 1577 -----LSGLQDPIIARMSICSEDKSPSECSLIASSPEESWPA----- 1614
Db 1669 PLISSPLKSVVSPKSRVDVSISSAKITWASS-----LSSPVKQMPGHAEVALVNGSISPLKY 1725
Qy 1615 COKAYNLNRTSTVTNNN-TAPTNRANQ-----NFEIEGIRETSQVI-----LRPGSPN 1665
Db 1726 ASSSTLINGCKATATLOEKISSATNSVSVSAATDTVEKFVSTTTAMPFSPPLRSYVSA 1785
Qy 1666 PTAQVENLKSMAHQRSSSYRLSKDASBELHAASSET 1705
Db 1786 PSFQSLRTPSAS-----ALYTSL-----GSSISATTSVT 1816

RESULT 14
S37771
ankyrin, erythrocyte - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S37771
R:Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an
A:Reference number: S37771; MUID:93252825; PMID:8486643
A:Accession: S37771
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1848

A:Cross-references: UNIPROT:Q61302; EMBL:X69063; NID:g311816; PIDN:CAA48801.i; PID:g3118
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:48-80/Domain: ankyrin repeat homology <AN01>
F:81-113/Domain: ankyrin repeat homology <AN02>
F:114-146/Domain: ankyrin repeat homology <AN03>
F:147-175/Domain: ankyrin repeat homology <AN04>
F:176-208/Domain: ankyrin repeat homology <AN05>

F:209-241/Domain: ankyrin repeat homology <AN06>
F:242-274/Domain: ankyrin repeat homology <AN07>
F:275-307/Domain: ankyrin repeat homology <AN08>
F:308-340/Domain: ankyrin repeat homology <AN09>
F:341-373/Domain: ankyrin repeat homology <AN10>
F:374-406/Domain: ankyrin repeat homology <AN11>
F:407-439/Domain: ankyrin repeat homology <AN12>
F:440-472/Domain: ankyrin repeat homology <AN13>
F:473-505/Domain: ankyrin repeat homology <AN14>
F:506-538/Domain: ankyrin repeat homology <AN15>
F:539-571/Domain: ankyrin repeat homology <AN16>
F:572-604/Domain: ankyrin repeat homology <AN17>
F:605-637/Domain: ankyrin repeat homology <AN18>
F:638-670/Domain: ankyrin repeat homology <AN19>
F:671-703/Domain: ankyrin repeat homology <AN20>
F:704-736/Domain: ankyrin repeat homology <AN21>
F:737-769/Domain: ankyrin repeat homology <AN22>
F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match 5.8%; Score 516.5; DB 2; Length 1848;
Best Local Similarity 19.6%; Pred. No. 1.2e-21;
Matches 397; Conservative 281; Mismatches 651; Indels 697; Gaps 81;
Qy 21 LKALLEKCKDYDERNECQGTPLMAAEOGNYEIVKELKNGANCNLEDLDNWTALISAK 80
Db 98 VRELNVYGANVVAQSQKGTPLYMAAQNHLEVKVFLLENGANQNVATEDGFTPLAVALQ 157
Qy 81 EGH----- 83
Db 158 QGHENVVAHLNYGTGKVRPLPALHIAARNDDTRTAAYLLQNDPNPDLVSKTGTPLHTA 217
Qy 84 -----IHVEILLKSGASLEHRDGMGTALMAWACVKGRTDVVELLSHGANSVVGLOYS 138
Db 218 AHYENLVAQULLNNGASVNFPPQGITPLHISRGVNIWVRLLLDRGAQIE-TRTKDE 276
Qy 139 VYPIWAAGRGHADIHLLLLQNGAKVNGSKYGTTPPLVMAARKHLECVGHLLAMGADV 198
Db 277 LTPLHCAARNGHVRISIELLDHGAPIQAKTKNGLSPIHMAAQGDHDCVRLLLQYNAEID 336
Qy 199 -----QEGA-----NSMTALIVAVKGGYTSQVKEILK 225
Db 337 DITDLHTPLHVAACHGHRVAKVLLDKGAKPNSRALNGFTPLHIACKNHIRVMELLEK 396
Qy 226 RNPVNLTKDGNLTALMASKEGHEIIVODLLDAGTYVNIPIRSGDTVLIGAVRGHVEI 285
Db 397 TGA SIDAVTESGLTPLHVASFNGHLPIVKNLLQKASPNVSNVKTETPLHVAARAGHTEV 456
Qy 286 VRALLQKYADIDIRGQDNKTALYWAKEGNAWVRDILQCNPDTEICTKOGTEPLIKATK 345
Db 457 AKYLLQNKAKANAKAKDDQTPHCAARIGHTGWVKLLLENGASPNLATTAGHTPLHTAAR 516
Qy 346 MRNIEVVELLLDKGAKVSAVDKKGDTPLHVAIKGRSRLLAELLLNPKDGLLYRPNKAG 405
Db 517 EGHVDTALALLEKEASQACMTKGTFTPLHVAARYKVKRLAEELLEHDA-----HPNAA 570
Qy 406 E---TPYNIDCSHQ-----KSIL-----TOIFGAR----- 427
Db 571 KNGLTPLHVAVHNNLIDIVKLLLPGRGSPHSPANNGYTPHLHIAKQNIQIIVARSLLQYGG 630
Qy 428 -----HLSPTETDGMGLYDLYSSALADILSEPTMQP-----PIC-- 462
Db 631 SANAESVQGVTPFLHAAQEGHTEMVALLLSKQANGNLGNKSGLTPLHLVLSQEGHVPVADV 690
Qy 463 -----VGLYAQWGSOK--SFLILKKLED-EMKTFACQQTPEPLPQFSWL 501
Db 691 LIKHGVTVDATFRMGYTPLHVASHYGNIKLVKFLLOHQADYNAKTKLGY--SPLHQAAQQ 748
Qy 502 --IVFLTLLCGGLGVFAFPVD-----TNLATAISLPLALIYIPFIVYFGRRGE 553
Db 749 GHTDIVTILLKNG-----ASPNEVSSNGTTPPLAKRGLYISVTDVLKVT----- 794
Qy 554 SSWMAWALSTRLARHIGYLELLFKLMFVNPPPELPEQTTKALPV---RFLFTDYNRLSS-- 608

795 --DTSVLVSKHRMSYPETVDEILDVS-----EDSGTAHISIMGDELVSGKAERRDSRD 848
609 VGETSLAEMIATLSDACEREFGLATRLFR-----VFRT-ESQKKKKWKTCCCL 658
849 VGEKELLDPVKLDQVVE-----SPAIPRIPCVTPETVIRSEDEQASKEYDESDLI 902
659 PSFVIFLIVGCIAGITLAILFRVDPKHLTVNAILISIASVGLAFVLCNCRTWQVLDLS 718
903 PS-----SPATETSONI-SPVASPVHTGFLVSF-----WVDA 933
719 LINSORKRLHSAASKLHLKXSEGMVKLKEVELMARMMAKTIDSFTQNTQRLVVIIDGLD 778
934 RCGSMGRSRN-----GLRVVIP-----PITCAAPRITCLV----- 966
779 ACQDKVQLMDTVRVLFSGKPFIAIPASDPHIIIIKAINONLSVLDRNSNINGHDYMRN- 837
967 -----KPKQLNTPPL-----AEEGLASRIIAGLPTGAQFLSPV 1001
838 IVHLPVPL-NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVQSNLSGEMTKLGSK 896
1002 IVEIPHASHGRG-----DRELVLVRSENGSV-----WKEHKSRYGESYLDQILN-GMD 1049
897 TALNRRTYRRRQMQRTITRQMSFDLTKLIVT---EDWFSDDISQPMRRLNIVSVTGR 953
1050 EELGSLEELEKRVCRIT--TDPLFYFVIMSRLCQD-YDTIGPEG-----GSLRSKL 1099
954 LRANQITFNWDRLASINLTQWPYRTSWILYLBETEGPLDQMTLKTMYERISKNIPTT 1013
1100 VPLVQATFPFENAVTNVKLALQ-----AQVPDELVTKLGNQ----- 1137
1014 KDVEPLEIDIGDIRNFEVFLSSRTPLVARDVKTFPLCTVNLDPKLEIIVADVRAAREQI 1073
1138 -----ATFSP-IVTVPRRRKFRPI----- 1157
1074 NIGSLAYPPLHEGPPRPSGYSQPAS---VCSSASFNPGFPFGVVPSPHSSYSYGLS 1130
1158 --GLRIPLPSPWTDNPR-DSGEGDTTSLRLCSVI-----GGTDQAQ-----WEDIT 1201
1131 GPQHIFY-NRAA-----VPATG-----SSILLSMVTVDVUCEKLRLQIEGLD 1170
1202 GTTKLIYANECANFTNVSARFWSLDCPRTAEAVHFATLLYKELTA----- 1247
1171 QNMMPQYCTTIKANIN---GRVLSQCNIDE-LKEMAMNFGDHLFRSMVLEM----- 1220
1248 ---VPYMAKFIIVAKMDAREGLRCYCMTDKDKVDTLEQHENFVEVARSDIEVLGMP 1304
1221 --RSVESQVVP-----EDPRFLNENSSAPVPHGESARR-----S 1252
1305 LFAELSGNLVPVKAQAQORGFHOSFRENRLAIPVKVRDSSRPBGFLSLFKTKMYEDT 1364
1253 SH-----TEPLTEL-----SOTPYTLNFSFEELNTLGLDEGAPR----- 1288
1365 QHILCHLNITMPPCTKGGAEDRRRTLTPLTRYLSILSESLRGLGTSDDTVENRMAVIRE 1424
1289 HSNLSWQSQR-----RTPSLSLNSQDSSI-----EISKLTDKVOAEY 1327
1425 HLGSLWELARELOFSVEDINRIEVENPNSLLDQSTALLIIVWDREGENAKM-ENLYTAL 1483
1328 RDAYREYIAQMSLEGSTISGRSPHSTHYIGSSSGGSIHSTLEOERGKEGELKQ 1387
1484 RNIDRSEIVNMLE-----VSGRQS-----RNLKPER----- 1509
1388 EDGRKSLMKRGDVIDYSS-----SGVSTNEASPLDPITEDEKSPQSGSKLLPGKSS 1441
1510 RHGREYSLSPSQVNGYSSLODELLSPASLOYP-SPLCADQVWNEVTVDAIPLAATE 1568
1442 ERPSLFOTDLKLKGGGLRYQKLPSEDESGTGRVQITPHCSKMIRTKRLKAKORECASPO 1501
1569 HDTMLENSDMQVMSAGLT-PSLVTAEDSS-----LECSKAE 1603
1502 EHSAPRTFTKAEYLSALLDKKSDSDGVRNENSSPNHSLHNEAADSQLEKANILE 1561
1604 D-----SDAIPWK-----LEGASDPTQGPGLDIVE 1633

QY 1562 LEDEGHSKRGMPHSISGLQDPIIARMSICSSEDKKSPSECSLI-----A 1605
Db 1634 DTVDSDATNGLADLLG--QQRVHARIT-----DSPSVRVQVLDRSQARTLWDKOGSTA 1685
QY 1606 SSPEE-----SW--PACQKAYNLNRTPTSVTLNNNTAPTNRANQNFDEI-----EGI 1650
Db 1686 VHPQEQATQSSQBEVTOGPHSFQRRITTI-----QGPEPGALQEQEVLVSTREHVQGRP 1740
QY 1651 RETSQVILRPGPSNPNTA-----VQENLKSMAHKRSQSSSYT 1688
Db 1741 PETGSP--KAGKEPSLWAPESAFSQEVOGDELQNPGEQVTEQFT 1784
RESULT 15
T42714
ankyrin 3, splice form 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42714
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42714
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <PEP>
A;Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Intons: 1387/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 5.8%; Score 510.5; DB 2; Length 1765;
Best Local Similarity 20.1%; Pred. No. 2.6e-21;
Matches 396; Conservative 278; Mismatches 641; Indels 653; Gaps 78;
QY 14 EENIPALKALLEKCKVDVERNECGQTPMLAAEQGNVEIVKELKNGANCNLEDLNDWT 73
Db 66 KEGHVEVSELLQREANVDAAATKKGNTALHIASLAGOAEVVKVLTNGANVNAQSONGT 125
QY 74 ALISASKEGHIHVEELLGSGASLEHRDMCGWTALMWACVKGRTDVVELL----- 124
Db 126 PLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQQQHDQVVSLELNDTKGVR 185
QY 125 -----SHGANPSVTGL----- 135
Db 186 LPAHHTAARKDDTKAAALLLQNDTNADVRSKSGFTPLHTAAHYGNINVATLLNRAAVID 245
QY 136 ---QYGVYPIIWAAGRHADIHLLLONGAKVNCSDKYGTTPPLVWAARK----- 182
Db 246 FTARNDDITPLHVASKGNAMVVKLLDRGAKIDAKTRDGLTPLHCGARSQHEQVEMLLD 305
QY 183 -----HLECVKHLAMGADVDEGANSMTALIVAKGVYQS 219
Db 306 RSAPILSKTKNGSLPLHMATQGDHLNCVQLLQHNVPVDDVTNDYLTALHVAACHGHYKV 365
QY 220 VKEILKRNPNVNLTDKGNLTALMASKE----- 247
Db 366 AKVLLDKASPNKALNGFTPLHIAACKKNRIRVWELLKHGASIQAVTESGLTPIHVAAF 425
QY 248 -GHIEIVQDILLDAGTVNIPDRSGDVTLVGAVGGHVEIVRALLQKVADIDIRQDNKTA 306
Db 426 MGHWNIVSQLMHHGASPNNTNVRGTALHMAARSQAEEVVRVYLVQDGAQVEAKKDDQTP 485
QY 307 LYVAVEKGNATMYRDIILQCNPDTEICTKGETPLIKATKMRNIEVVVELLDKGAQVSAVD 366
Db 486 LHISARLGRADIVQQLLQOQASPNAAATSGTTPHLHAAAREGHEDVAFLDHGASLITT 545

QY	367	KKGDTPLHVAIRGRSRLAEILLR : : : :	-----NPKDGRLL-----	398
Db	546	KKGTPTLHVAAYKGLVASLLLOKSASPDAGKSGSLTPLHVAAHYNOKVALLLLDOGA : : : :	-----NPKDGRLL-----	605
QY	399	--YRPNKAGETPYNIDC--SHOKSILTQL--FCAR-- : : : :	-----HLSPTETDGDMLGY 441	
Db	606	SPHAANKNGYTPHLIAAKKNQMDATSLLEVGADANAVTROGIIASVHLAAOGEHVDMVSL : : : :	-----HLSPTETDGDMLGY 441	665
QY	442	DLYSSALADIUSEPTMOP : : : :	-----PICVGLYAQM 469	
Db	666	LLSNANVNLSNKSGLTPLHLAAQEDRVNAEVLNQGAHVDAQTKMGYTPLHWGCH--Y 723 : : : :	-----PICVGLYAQM 469	723
QY	470	GSGK--SFLLK--KLEDEMTFAGQOTEPILFOFSWLIIVFLTL : : : :	-----LLCGGLGV 516	
Db	724	GNIKIVNFLQHSKAVNAKTNGYTAHQAAAQQGHTTHIINVLQNNASPNELTVNG : : : :	-----LLCGGLGV 516	779
QY	517	FAPVDVTNLAIASISFLALIYIFPVI : : : :	-----PICVGLYAQM 469	544
Db	780	-----NTALAIAARRLGYSISVDLTALKVTEEMTWTTITEKHKNVPETMNEVLMSDDEV 834 : : : :	-----PICVGLYAQM 469	834
QY	545	-----YPGRRGESNNAWALSTRLARHIGYLELPKL : : : :	-----MPVN 582	
Db	835	RKASAPEKLSDEYISDCEEGDKCTFKIPKVQ-- : : : :	-----EVLVKSEDAITGTDKYLG 886	
QY	583	PPELPEQTTRALPVRFETDYNRLSSVGGETSIAEMIATLSDACEREFGLATLRFVRFR 642 : : : :	-----EVLVKSEDAITGTDKYLG 886	642
Db	887	PODLKEGGDSLPAE-----GYVG--FSLGARSASUR-----SPSSDRSYTLNR 928 : : : :	-----SPSSDRSYTLNR 928	928
QY	643	TEESQGKKWKKTCCLPSSFVIFLVGCIITALLAIFRVDP--KHLT-----VNAI 693 : : : :	-----VNAI 693	693
Db	929	SSYAR-----DSMWIEELLVPSKEQHLLFTREFDSDSURHYSWAADTLDNVNV 977 : : : :	-----VNAI 693	977
QY	694	LISIASVVGLAFVLCRTWMQVLSLNSQ-----RK-----RLHSAASKLHL- 737 : : : :	-----RLHSAASKLHL- 737	737
Db	978	SSPVHSGFLVSFMVDARG-----GSMRGSRHGMRIIIPPRKCTAPTTRITCRLVKRUKLA 1032 : : : :	-----RLHSAASKLHL- 737	1032
QY	738	-----KSEG-----FMKVLKCEVELMARMAKTIDSSTQN 766 : : : :	-----FMKVLKCEVELMARMAKTIDSSTQN 766	766
Db	1033	NPPPWVEGEGLASRLVEMGPAGAFLGPVIVEIPHFGSMRGKERELIVLRSENGETWEH 1092 : : : :	-----FMKVLKCEVELMARMAKTIDSSTQN 766	1092
QY	767	Q-----TRLVVIIDGLDACEODKVLQMLDTRV--LFSGKGPFAIFASDPHHIIKAINQN 819 : : : :	-----FMKVLKCEVELMARMAKTIDSSTQN 766	819
Db	1093	QFDSKNEDLAEELLGMD--EELDSPSELGKTCRIITK--DFPQYFA-----VVSRIKQE 1144 : : : :	-----VVSRIKQE 1144	1144
QY	820	LNSVLRDSNTINGHDYMRNI VHLPVFLNSRGLSNARKFLV-TSANGDITCSDDTTGTQ--- 875 : : : :	-----VVSRIKQE 1144	875
Db	1145	SNOIGPBGGI-----LSSTVPLVQAASFPEGALTCKRIRVGLQAQP 1184 : : : :	-----VVSRIKQE 1144	1184
QY	876	--EDTDERVSONSIGENTKLGSKTALNRDTY--RRQMORTITROMSFDLTCLL---VT 928 : : : :	-----VVSRIKQE 1144	928
Db	1185	VPETVVKI-----LGNKATFPVIVPEPRRRKFKPIIT--MTIPVPPPGSGEVGS 1232 : : : :	-----VVSRIKQE 1144	1232
QY	929	EDWFSDISPOTMRLLNIVSVTG-----RLIRANQITFNMDRLASNINLTQWPVRT 980 : : : :	-----VVSRIKQE 1144	980
Db	1233	NGYKGDATPNL--RL--CSITGTSQAOWEDITGTTPLFIKDVCVFVTVNSARF---- 1284 : : : :	-----VVSRIKQE 1144	1284
QY	981	SWL-----ILYBETEGLPOMTLKTYMERISKNIPTTKOVPELLEDGDIRNEFEVLSS 1035 : : : :	-----VVSRIKQE 1144	1035
Db	1285	WLADCHQVL---ETVGLASQLYRELIC-----VPY-----MAKFVVFPAKT 1321 : : : :	-----VVSRIKQE 1144	1321
QY	1036	RTPLVARDVKTFLPCTVNLNDPKLREIIADVRAREQINIGGLA-YPLLPHGEGPPRPPS 1094 : : : :	-----VVSRIKQE 1144	1094
Db	1322	NDP-----VESSLRCFCMTDDRVDKTL-----EQENEEVARSKDIEVLEGKPIYVD 1369 : : : :	-----VVSRIKQE 1144	1369
QY	1095	GYSQPASVCSASSFNGFPFGVSPQPHSSYYISGLSGFHPP-----YNRAAVPATGSSLL 1150 : : : :	-----VVSRIKQE 1144	1150
Db	1370	CYGNLAPLTK-----GG-----QQLVNFYS-FKENRLPFISKIRDTSQPCGRLSFL 1416 : : : :	-----VVSRIKQE 1144	1416
QY	1151	LSMTVDVCEKURLQIBGLQNMMPQVCTTIKANINGRVLSCQNDIELKKEWANPFGDW 1210 : : : :	-----VVSRIKQE 1144	1210
Db	1417	KEPKTT-----KGLPQTAV-----CNLMITIPAHKKAFAKDRROS----- 1451 : : : :	-----VVSRIKQE 1144	1451

Qy	1211	HLFRSWLENR-----SVESQVVPDPRLPNNSGAPVPHGESARRSSHTELPFLTELSQTP	1367
Db	1452	--FASLALRKYSYLTEPSPMSPOCERTDIRMAIV-----ADHLGLSWTELARE--	1499
Qy	1268	YTLNFPEELNTLGLDEGAPRHNSL-----W--QSQTRTPSLGSLNSQDSSIEI	1316
Db	1500	--LNFVDEINQRVEN---PNSLISQSFMLLKKWTRDGNATTDALTSLVTKINRIDI	1554
Qy	1317	SKLTDKVOAERYDAYREYIAQMSQLEGGTGSSTISGRSPHSTYYIG--QSSSGGSIHST	1374
Db	1555	VTLLEGPIDYGNI-----SCTRFSADENNVPDVGWQNETSGSLSP	1600
Qy	1375	LEOERKEGELKQEGCRKFLMKRGVDIVSSGSGVTNEASPLDPTTEBDEKSDQSG---	1431
Db	1601	AQARLLTGGLLDRD-----DSSQARDISITY-----LTGPSPGKTEANGNHT	1643
Qy	1432	SKLLPGKKSERPSLFQDCLKGGGLRYOKLPDSDESGTGRVQIPTHCSKMIRTKRLK	1491
Db	1644	AEVPEAKA-----PYPPESQNDIGKQGIK-----ENLK	1673
Qy	1492	AKQREKASQEHSAEBIRTFIKAKEYLSDALD--KKDSSDGSGRVSNESPNHSLHNEAA	1549
Db	1674	PKTHGCGRTEE--PVSPLTAYQKSLEETSXLIVIEDAPKPCVPVGMK-----KMTRTTA	1724
Qy	1550	DDSQLEKANLIELEDGSHGKRGMPHSLGSLQDPIIARMSICSEDKKS	1597
Db	1725	DG----KARLNLOEBEGST--RSEPPQGEYK--VTKKEIRNVEKKT	1764

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:17:40 ; Search time 188.5 Seconds
(without alignments)
4658.970 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASBESTGFERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8853	100.0	1715	2 Q9ERD4	Q9erd4 rattus norv
2	8639.5	97.6	1762	2 Q9EQG6	Q9eqg6 rattus norv
3	8212	92.8	1777	2 Q9ULH0	Q9ulh0 homo sapien
4	7892	89.1	1693	2 Q80TG7	Q80tg7 mus musculu
5	7693	86.9	1554	2 Q80SX9	Q80sx9 mus musculu
6	6233.5	70.4	1680	2 Q7T163	Q7t163 brachydanio
7	6219.5	70.3	1672	2 Q6P7V1	Q6p7y1 brachydanio
8	5073.5	57.3	1031	2 Q9UF42	Q9uf42 homo sapien
9	3097	35.0	691	2 Q7Z322	Q7z322 homo sapien
10	2732	30.9	664	2 Q6MZU2	Q6mzu2 homo sapien
11	2421.5	27.4	1604	2 Q7KVP5	Q7kvp5 drosophila
12	2421.5	27.4	1626	2 Q7KVP6	Q7kvp6 drosophila
13	2421.5	27.4	1678	2 Q9W210	Q9w210 drosophila
14	2408	27.2	543	2 Q9H9X4	Q9h9e4 homo sapien
15	2329.5	26.3	1354	2 Q7PZV2	Q7pzy2 anopheles g
16	1876	21.2	1089	2 Q7YU92	Q7yu92 drosophila
17	1848.5	20.9	1398	2 Q20109	Q20109 caenorhabdi
18	1097	12.4	239	2 Q8K0F2	Q8k0p2 mus musculu
19	1024	11.6	246	2 Q9NT37	Q9nt37 homo sapien
20	706	8.0	257	2 Q7OMF9	Q7gmf9 anopheles g
21	599.5	6.8	129	2 Q9H899	Q9h899 homo sapien
22	556	6.3	1880	1 ANK1_HUMAN	P46157 homo sapien
23	555	6.3	426	2 Q8YTG9	Q8ytg9 anabaena sp
24	555	6.3	1856	2 Q99407	Q99407 homo sapien
25	551	6.2	1943	2 Q61307	Q61307 mus musculu
26	550	6.2	3924	1 ANK2_HUMAN	O01484 homo sapien
27	545	6.2	1719	2 Q13768	Q13768 homo sapien
28	543	6.1	1887	2 Q7Z3G4	Q7z3g4 homo sapien
29	531.5	6.0	1280	2 Q6N064	Q6n064 homo sapien
30	531.5	6.0	1863	2 Q7Z315	Q7z315 homo sapien
31	529	6.0	2622	2 Q70511	Q70511 rattus norv

32	527.5	6.0	1862	1 ANK1_MOUSE	Q02357 mus musculu
33	526	5.9	4377	1 ANK3_HUMAN	Q12955 homo sapien
34	523.5	5.9	1375	2 Q6ZSE6	Q6zse6 homo sapien
35	516.5	5.8	1848	2 Q61302	Q61302 mus musculu
36	510	5.8	1726	2 Q8VC68	Q8vc68 mus musculu
37	509.5	5.8	2206	2 Q7QAV8	Q7qav8 anopheles g
38	505.5	5.7	1136	2 Q9N180	Q9n180 bos taurus
39	504.5	5.7	1219	2 Q8C8R3	Q8c8r3 mus musculu
40	503.5	5.7	843	2 P97582	P97582 rattus norv
41	500.5	5.7	1486	2 Q8TEF1	Q8tef1 homo sapien
42	499.5	5.6	1159	2 Q9NCF8	Q9ncp8 drosophila
43	499.5	5.6	1571	2 Q7KU92	Q7ku92 drosophila
44	498	5.6	792	2 Q7Q172	Q7q172 anopheles g
45	493.5	5.6	830	2 Q96I86	Q96i86 homo sapien

ALIGNMENTS

RESULT 1
Q9ERD4
ID Q9ERD4 PRELIMINARY; PRT; 1715 AA.
AC Q9ERD4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ankyrin repeat-rich membrane-spanning protein.
GN Name=ARMS;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20585245; PubMed=11150334;
RA Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;
RT "An evolutionarily conserved transmembrane protein that is a novel
RT downstream target of neurotrophin and ephrin receptors.";
RL J. Neurosci. 21:176-185(2001).
DR EMBL; AF313464; AAC34167.1; -.
DR HSSP; P09959; 1SW6.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS00088; ANK_REPEAT; 10.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 1715 AA; 190532 MW; 505593C4A19A6CDC CRC64;

Query Match	100.0%;	Score	8853;	DB	2;	Length	1715;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1715;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MSVLISQSVINYVEENIPALKALLERCKVDNERNECQTPPLMLAAEQGNVEIVKELKN	60				
Db	1	MSVLISQSVINYVEENIPALKALLERCKVDNERNECQTPPLMLAAEQGNVEIVKELKN	60				
Qy	61	GANCNLEDLDNWTALISASKEGHIHIVEELIKSGASLEHRDMGGTALMWACYKGRDGV	120				
Db	61	GANCNLEDLDNWTALISASKEGHIHIVEELIKSGASLEHRDMGGTALMWACYKGRDGV	120				
Qy	121	ELLISHGANPSVTGLQYSVPIIWAAGRGHADIIVHLLQNGAKVNCSDKYTTPLVWAAR	180				
Db	121	ELLISHGANPSVTGLQYSVPIIWAAGRGHADIIVHLLQNGAKVNCSDKYTTPLVWAAR	180				
Qy	181	KGHLECVKHLAMGADVDDQGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDGN	240				
Db	181	KGHLECVKHLAMGADVDDQGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDGN	240				
Qy	241	LMTASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	300				
Db	241	LMTASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	300				

Db 241 LMIASKEGHIEIYQDLDLAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
QY 301 QDNKTALYWAVEKGNATWRDILQCPDPTIEICTKDGETPLIKATKRNIEVVELLDKGA 360
Db 301 QDNKTALYWAVEKGNATWRDILQCPDPTIEICTKDGETPLIKATKRNIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420
Db 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420
QY 421 TQIFGARHLSPETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKKL 480
Db 421 TQIFGARHLSPETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKKL 480
QY 481 EDEMKTFAQOQTPLQFQSWLIYVFLILLCCGGLGVFAPVPTDNLAIASLSFLALIYIF 540
Db 481 EDEMKTFAQOQTPLQFQSWLIYVFLILLCCGGLGVFAPVPTDNLAIASLSFLALIYIF 540
QY 541 FIVIFGGRREGESWNAWALSRLARHIGYLLPKLMPVNPPELPEOTTKALPVRELF 600
Db 541 FIVIFGGRREGESWNAWALSRLARHIGYLLPKLMPVNPPELPEOTTKALPVRELF 600
QY 601 TDYNRSLSSVGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQGGKKWKTCCLPS 660
Db 601 TDYNRSLSSVGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQGGKKWKTCCLPS 660
QY 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVLCNCTWQVLDLSLL 720
Db 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVLCNCTWQVLDLSLL 720
QY 721 NSORKRLHSAASKLHKLSGFGMKVLKCEVELMARMAKTIIDSTFQNTQTRLVWIIDGLDAC 780
Db 721 NSORKRLHSAASKLHKLSGFGMKVLKCEVELMARMAKTIIDSTFQNTQTRLVWIIDGLDAC 780
QY 781 EQDKVLQMLDTRVLFSGKPFIIAFASDPHIIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Db 781 EQDKVLQMLDTRVLFSGKPFIIAFASDPHIIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
QY 841 LPVFLNSRGLSNARKELVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKELVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
QY 901 RRDYTRRRQWRITTRQMSFDLTKLAVTEDFWSDISPTQWRRLINLVSVTGRLLRANQIT 960
Db 901 RRDYTRRRQWRITTRQMSFDLTKLAVTEDFWSDISPTQWRRLINLVSVTGRLLRANQIT 960
QY 961 FNWDLASWINLTEQWPYRTSWLILYLEETEGPLDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWINLTEQWPYRTSWLILYLEETEGPLDQMTLKTMYERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIADVRAAREQINIGGLAY 1080
QY 1081 PPLPLHEGPPRPGSGVQSPASVCSASFNGPFGVGVSPHSHSYTSGLSGPOHPFNRA 1140
Db 1081 PPLPLHEGPPRPGSGVQSPASVCSASFNGPFGVGVSPHSHSYTSGLSGPOHPFNRA 1140
QY 1141 AVPATGSSLLSSMTVDVCEKLRQIBGLDQNMWPOYCTTIKKANINGRVLSCNIDELK 1200
Db 1141 AVPATGSSLLSSMTVDVCEKLRQIBGLDQNMWPOYCTTIKKANINGRVLSCNIDELK 1200
QY 1201 KEMAMNFGDWHLFRSWVLENRSVESQVVPDPRFLNENSSAPVPHGESARRSHTEPLT 1260
Db 1201 KEMAMNFGDWHLFRSWVLENRSVESQVVPDPRFLNENSSAPVPHGESARRSHTEPLT 1260
QY 1261 ELSSQTPYTLNFPSEELNTLGLDEGAPRHSNLSWQSOTRTPPLSLNSQDSSIEISKLT 1320
Db 1261 ELSSQTPYTLNFPSEELNTLGLDEGAPRHSNLSWQSOTRTPPLSLNSQDSSIEISKLT 1320
QY 1321 DKVQAEYRDYREYIAQMSQLEGGTSGSTISGRSSPHSTYIYIQGSSGGSIHSTLEOERG 1380
Db 1321 DKVQAEYRDYREYIAQMSQLEGGTSGSTISGRSSPHSTYIYIQGSSGGSIHSTLEOERG 1380

QY 1381 KBGELKQEDGRKSFMLMRGDVIDYSSGYSTNEASPLDPTITEDEKSDOSGSKLLPGKKS 1440
Db 1381 KBGELKQEDGRKSFMLMRGDVIDYSSGYSTNEASPLDPTITEDEKSDOSGSKLLPGKKS 1440
QY 1441 SERPSLFQTDLKLKGGGLYQKLPSEDESGTGRVQITPHCSKMIRTKLKAQORECASP 1500
Db 1441 SERPSLFQTDLKLKGGGLYQKLPSEDESGTGRVQITPHCSKMIRTKLKAQORECASP 1500
QY 1501 QHSABPIPTFKAKYILSDALLDKDSDSDGVRNNESSPNHSLHNEAADDSOLEKANLI 1560
Db 1501 QHSABPIPTFKAKYILSDALLDKDSDSDGVRNNESSPNHSLHNEAADDSOLEKANLI 1560
QY 1561 ELEDEGHSGKRGMPHSLGLQDPIIARMSICSEDKSPSECSLIASSPESWPACOKAYN 1620
Db 1561 ELEDEGHSGKRGMPHSLGLQDPIIARMSICSEDKSPSECSLIASSPESWPACOKAYN 1620
QY 1621 LNRTPTSTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPGSPNPTAVQNLKSMHAK 1680
Db 1621 LNRTPTSTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPGSPNPTAVQNLKSMHAK 1680
QY 1681 RSQRSSYTRLSDASELHHAASSESTGFGEERESIL 1715
Db 1681 RSQRSSYTRLSDASELHHAASSESTGFGEERESIL 1715

RESULT 2
Q9EQG6 PRELIMINARY; PRT; 1762 AA.
ID Q9EQG6; AC Q9EQG6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIDINS220.
GN Name=Kidins220;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568256; PubMed=10998417; DOI=10.1074/jbc.M005261200;
RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozengurt E.,
RA Schiavo G.;
RT "Identification and cloning of Kidins220, a novel neuronal substrate
of protein kinase D.";
RL J. Biol. Chem. 275:40048-40056(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Mitchell M., Schiavo G.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF239045; AAG35185.2; -.
DR HSSP; P09959; 1SW6.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 10.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1762 AA; 195714 MW; 0CB2689A571F8AE4 CRC64;

Query Match 97.6%; Score 8639.5; DB 2; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;
QY 1 MSVLISQSVINYVEENIPALKLEKCKVDVERNECGQTPLMLAAEQGNVEIVKELKN 60
Db 1 MSVLISQSVINYVEENIPALKLEKCKVDVERNECGQTPLMLAAEQGNVEIVKELKN 60
QY 61 GANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLDNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120

SQ	SEQUENCE	1777 AA; 197209 MW; B6505923FB45F143 CRC64;	
	Query Match	92.8%; Score 8212; DB 2; Length 1777;	
	Best Local Similarity	89.7%; Pred. No. 0;	
	Matches 1590; Conservative	53; Mismatches 70; Indels 60; Gaps 4;	
Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAABQGNVEIVKELKN	60
Db	7	MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAABQGNLEIVKELIKN	66
Qy	61	GANCNLEDLNTWLTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV	120
Db	67	GANCNLEDLNTWLTALISASKEGHVHIVEELLKCGVNLHRDMGWTALMWACYKGRDVV	126
Qy	121	ELLISHGANPSVTGLQSVVPTIWAAGRGHADIVHLLQLONGAKVNCSDKYGTTPLVWAAR	180
Db	127	ELLISHGANPSVTGL-QSVVPTIWAAGRGHADIVHLLQLONGAKVNCSDKYGTTPLVWAAR	185
Qy	181	KHLECVKHLLANGADVQEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA	240
Db	186	KHLECVKHLLANGADVQEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA	245
Qy	241	LMIASKEGHIIEVQDILLDAGTYNIPDRSGDTVLIGAVRGHVEIVRAALQKYADIDIRG	300
Db	246	LMIASKEGHIIEVQDILLDAGTYNIPDRSGDTVLIGAVRGHVEIVRAALQKYADIDIRG	305
Qy	301	QDNKTALYWAVERKGNATWDRDILQCNPDTEI CTYKGETPLIKATKORNIIEVVELLDKGA	360
Db	306	QDNKTALYWAVERKGNATWDRDILQCNPDTEI CTYKGETPLIKATKORNIIEVVELLDKGA	365
Qy	361	KVSADVKKGTPLHVALRGSRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL	420
Db	366	KVSADVKKGTPLHVALRGSRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL	425
Qy	421	TOIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSFKLKKL	480
Db	426	TOIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSFKLKKL	485
Qy	481	EDEKTFAGQOTPELFOFSLIIVFLTLILCGGLVPAFPVDNLTALAIISFLALIYIF	540
Db	486	EDEKTFAGQOIPELFOFSLIIVFLTLILCGGLVPAFTVHPNLGIAVSLFALIYIF	545
Qy	541	FIVYFGREGESWNWALSTRLARHIGVLELLPKLMFVNPPPELPEQTTKALPVRFIF	600
Db	546	FIVYFGREGESWNWALSTRLARHIGVLELLPKLMFVNPPPELPEQTTKALPVRFIF	605
Qy	601	TDYNRLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFRTEESQGGKKWKTCCLPS	660
Db	606	TDYNRLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTCGKKKKWKTCCLPS	665
Qy	661	FVIFLFIVGCIIGITILLAIIFRVDPKHLTVNAILISTASVVGAFVNLCTWQVLDLIL	720
Db	666	FVIFLFITIGCIISGITILLAIIFRVDPKHLTVNAILISTASVVGAFVNLCTWQVLDLIL	725
Qy	721	NSQKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTTIDSTFQNTQRLVLIIDGLDAC	780
Db	726	NSQKRLHNAASKLHLKSGFMKVLKCEVELMARMAKTTIDSTFQNTQRLVLIIDGLDAC	785
Qy	781	EQDKVLQMLDTRVLFVSKGPPIAIFASDPHIIKAINONLNSVLDRSNGINGHDMRNIVH	840
Db	786	EQDKVLQMLDTRVLFVSKGPPIAIFASDPHIIKAINONLNSVLDRSNGINGHDMRNIVH	845
Qy	841	LPVFLNSRGI SNARKFLVTSATNGDITCSDTTGTQEDTDRVQNSIGEMTKLGSKTALN	900
Db	846	LPVFLNSRGI SNARKFLVTSATNGDVPDSDTTGTQEDADRRVQNSIGEMTKLGSKTALN	905
Qy	901	RRDYRRRQORITTRQMSFDLTKLIVTEDWFSDISPQTMRRLLINIVSVTGRLLRANQIT	960
Db	906	RRDYRRRQORITTRQMSFDLTKLIVTEDWFSDISPQTMRRLLINIVSVTGRLLRANQIS	965
Qy	961	FNWDLASWINLTQWQPYRTSWLILYLEETEGIPDQWTLKTYVERISKNIPTTKQVBEPLL	1020
Db	966	FNWDLASWINLTQWQPYRTSWLILYLEETEGIPDQWTLKTYVERISKNIPTTKQVBEPLL	1025

Qy	1021	EIDGDIRNEFEVLSSTPVLVARDVKTFELPCTVNLDPKLEIITADYRAAREQINIGLAY	1080
Db	1026	EIDGDIRNEFEVLSSTPVLVARDVKTFELPCTVNLDPKLEIITADYRAAREQISIGLAY	1085
Qy	1081	PPLPLEHEGPRPSPGYSQSPASVCSSASFNPGPPGGVVSQPHSSYYSGLSGPQHPFVN--	1138
Db	1086	PPLPLEHEGPRAPSPGYSQSPASVCSSTSFNPGPAGGVVSQPHSSYYSGMTGPPHPFNR	1145
Qy	1139	-----RAAPV	1143
Db	1146	FFAPYLYTPRYPPGSGHILSRPSVKTSLPRDQNGLEVIKEDAEAGLSPTDSRSGSP	1205
Qy	1144	ATGSSLLLSMTVDVCEKLEIIEGLDQNMPOYCTTIKKANINGVLSQCNIDELKKEM	1203
Db	1206	APGFVLLNSLNDVACEKLEIIEGLDQNMPOYCTTIKKANINGVLSQCNIDELKKEM	1265
Qy	1204	AMNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENESAPVPHGESARRSHTLPLTELS	1263
Db	1266	NMFGDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHTELS	1325
Qy	1264	SQPTYLNFSPFELNLTGLDEGAPRHSNLWSQOTRRTPSLSLNSQDSIESIKLTDKV	1323
Db	1326	SQPTYLNFSPFELNLTGLDEGAPRHSNLWSQOTRRTPSLSLNSQDSIESIKLTDKV	1385
Qy	1324	QAEYRDAYREYTAQMSOLEGGTSGSTISGRSSPHSTYVIGOSSSGSIHSTLEQERKEG	1383
Db	1386	QAEYRDAYREYTAQMSOLEGGTSGSTISGRSSPHSTYVIGOSSSGSIHSTLEQERKEG	1445
Qy	1384	ELKQEDGRKSFMLKMGDVIDYSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKSSER	1443
Db	1446	EPKPDGRKSFMLKMGDVIDYSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKSSER	1505
Qy	1444	PSLQFQDQLKKGGLRYQKLPDEDESGTCRVOITPHCSKMIRTKLKAQORCAGPOEH	1503
Db	1506	SSLQFQDQLKKGGLRYQKLPDEDESGTCESDNTF-LLKDDKDKRAEGKVERVPKSPKH	1564
Qy	1504	SAEPIRTFTKAKBYLSDALLDKKSDSGVRNESPNSHNEAADDQLEKANIIELE	1563
Db	1565	SAEPIRTFTKAKBYLSDALLDKKSDSGVRNESPNSHNEAADDQLEKANIIELE	1624
Qy	1564	DEHSGKRGMPHSLGSLQDPIIARMSICSEDKSPSECSLIASSPESWPACQAYNLNR	1623
Db	1625	DDSHSGKRGMPHSLGSLQDPIIARMSICSEDKSPSECSLIASSPESWPACQAYNLNR	1684
Qy	1624	TPSTVTLNNTAPTNRANONFDEIGRETQSIVILRPGSPNPNTAVONENLKSMAHKRSQ	1683
Db	1685	TPSTVTLNNTAPTNRANONFDEIGRETQSIVILRPGSPNPNTAVONENLKSMAHKRSQ	1744
Qy	1684	RSSYTRLSKXDASELH-AASSESTGFEERESIL	1715
Db	1745	RSSYTRLSKXDASELH-AASSESTGFEERESIL	1777

RESULT 4

Q80TG7	PRELIMINARY;	PRT; 1693 AA.
ID	Q80TG7	Q80TG7;
AC	Q80TG7	01-JUN-2003 (T-EMBLrel. 24, Created)
DT	01-JUN-2003	(T-EMBLrel. 24, Last sequence update)
DT	01-MAR-2004	(T-EMBLrel. 26, Last annotation update)
DE	MKIAA1250	protein (Fragment).
GN	Name=C330002119Rik; Synonyms=mkIAA1250;	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Brain;	
RX	MEDLINE=22579291; PubMed=12693553;	
RA	Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,	
RA	Nakajima D., Nagase T., Ohara O., Koga H.;	

RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 DR EMBL; AK122478; BAC65760.1; --
 DR HSSP; P09959; 18W6.
 DR MGD; MGI:1924730; C330002119Rik.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 10.
 DR PRINTS; PR01415; ANKTRIN.
 DR SMART; SM00248; ANK; 9.
 DR PROSITE; PS50088; ANK_REPEAT; 8.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR KW ANK repeat.
 FT NON TER
 SQ SEQUENCE 1 1693 AA; 188253 MW; 599DE2A3419D4C9D CRC64;
 Query Match 89.1%; Score 7892; DB 2; Length 1693;
 Best Local Similarity 90.1%; Pred. No. 0;
 Matches 1544; Conservative 30; Mismatches 39; Indels 100; Gaps 6;
 QY 83 HIHIVEELLKSGASLEHRDMCGWTALMWACYKGRDVTVELLSHGANSPTVGLQYSVYPI 142
 DB 1 HIHIVEELLKCGANLEHRDMCGWTALMWACYKGRDVTVELLSHGANSPTVGL-QYSVYPI 59
 QY 143 IWAAGRGHADIVHLLQNGAKVNSDKYGTTPVWAARKHLECKVHLLANGADVDEGA 202
 DB 60 IWAAGRGHADIVHLLQNGAKVNSDKYGTTPVWAARKHLECKVHLLANGADVDEGA 119
 QY 203 NSMTALIVAVKGGYTQSVKEILKENPNVNLTKDGNALMIASKEGHEIEIVQDILLDAGTY 262
 DB 120 NSMTALIVAVKGGYTQSVKEILKENPNVNLTKDGNALMIASKEGHEIEIVQDILLDAGTY 179
 QY 263 VNIIPDRGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTALYVAVEKGNATWVRDI 322
 DB 180 VNIIPDRGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTALYVAVEKGNATWVRDI 239
 QY 323 LQCPNPTETICTKGETPLIKATKRNIEVVELLDKAKYSAVDKKGDTPLHVAIRGRSR 382
 DB 240 LQCPNPTETICTKGETPLIKATKRNIEVVELLDKAKYSAVDKKGDTPLHVAIRGRSR 299
 QY 383 RLAEILLRNPKDGLLYRPNKAGETPNIDCSHQKSLITQIFGARHLSPTETDGMGLGYD 442
 DB 300 RLAEILLRNPKDGLLYRPNKAGETPNIDCSHQKSLITQIFGARHLSPTETDGMGLGYD 359
 QY 443 LYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKLEDEDMKTFAGQQTPEPLPQFSWLI 502
 DB 360 LYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKLEDEDMKTFAGQQTPEPLPQFSWLI 419
 QY 503 VFLTLCCGGLGVFAPVDNLAIAISLSFLALYIFVIFYGGRREGESNNWAWALS 562
 DB 420 VFLTLCCGGLGVFAPVDNLAIAISLSFLALYIFVIFYGGRREGESNNWAWALS 479
 QY 563 TRLARHIGYLELLEPKLMFVNPPPEQTTKALPVRFLETDVNRSLSSVGGETSIAEMIATL 622
 DB 480 TRLARHIGYLELLEPKLMFVNPPPEQTTKALPVRFLETDVNRSLSSVGGETSIAEMI --- 536
 QY 623 SDACERFGLFATLRFVFRTEBSQGGKKWKTCCLPSPVIFLPIVGCIIAGITLLAIFR 682
 DB 537 -----FEVLRLLKILRV-----KRNKGKTCCLPSPVIFLPIVGCIIAGITLLAIFR 581
 QY 583 VDPKHLTVNAILISIASVGLAFVLCRTWQVLDLSLNSQKRLHSAASKHLKXSEGF 742
 DB 582 VDPKHLTVNAILISIASVGLAFVLCRTWQVLDLSLNSQKRLHSAASKHLKXSEGF 641
 QY 743 MKVLKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDACEQDKVQLMDTVRVLFSGKPGFI 802
 DB 642 MKVLKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDACEQDKVQLMDTVRVLFSGKPGFI 701
 QY 803 AIFASDPHIIKAINQNLSVLRDSNINGHDMYMRNIVHLPVFLNSRGLSNARKFLVTSAT 862
 DB 702 AIFASDPHIIKAINQNLSVLRDSNINGHDMYMRNIVHLPVFLNSRGLSNARKFLVTSAT 761
 QY 863 NGDITCSDTTCQEDTDRRVSONSLGEMTKLGSXTALNRDITYRRROMORTITQMSFDL 922
 DB 762 NGDISCSEATGVEDADRVSQNSLGEMTKLGSXTALNRDITYRRROMORTITQMSFDL 821
 QY 923 TKLJLVTDWFDSDISFQPMRRLNIVSVTGRLLRANQITFNWDRLASWNLNTEQWPYRTSW 982
 DB 822 TKLJLVTDWFDSDISFQPMRRLNIVSVTGRLLRANQITFNWDRLASWNLNTEQWPYRTSW 881
 QY 983 LILYLETEGLPDQMTLTKMYERISKNIPTTKQVEPLEIDGDIRNPFVFSRTPVLVA 1042
 DB 882 LILYLETEGLPDQMTLTKMYERISKNIPTTKQVEPLEIDGDIRNPFVFSRTPVLVA 941
 QY 1043 RDVKTLPCTVNLDPKLEIIVADVRAAREQINIGLAYPPLPLHEGPPRPSGYSQASV 1102
 DB 942 RDVKTLPCTVNLDPKLEIIVADVRAAREQINIGLAYPPLPLHEGPPRPSGYSQASV 1001
 QY 1103 C-SSASFNPGPPGGVSPQPHSSYISGLSGPQHPFYNR----- 1139
 DB 1002 C-SSASFNPGPPGGVSPQPHSSYISGLSGPQHPFYNRPPFAPVLYTPRYYPGSGHLIS 1061
 QY 1140 -----AAV 1142
 DB 1062 RSVKTSILPRDQNNGLVIEKEDAAEGLSPSTASSREKSWTRKQLMELCDSGFNKQOQASV 1121
 QY 1143 PATGSSLLLSMTVDVVCCKLROIEGLDONMMPQCTTIKKANINGRVLSCNIDELKKE 1202
 DB 1122 PATGSSLLLSMTVDVVCCKLROIEGLDONMMPQCTTIKKANINGRVLSCNIDELKKE 1181
 QY 1203 MAMNFGDHLFRSNVLEMRSVESQVPEPRFLNENSSAPVPHGESARRSHTLPLTEL 1262
 DB 1182 MAMNFGDHLFRSNVLEMRSVENQVPEPRFLNENSSAPVPHGESARRSHTLPLTEL 1241
 QY 1263 SSQTPYTLNPSFELNLTGLDEGAPRHSNLSWQSTRTTSLSSINSQDSSIEISKLTKDK 1322
 DB 1242 SSQTPYTLNPSFELNLTGLDEGAPRHSNLSWQSTRTTSLSSINSQDSSIEISKLTKDK 1301
 QY 1323 VQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYIYGSSSGGSIHSTLQERCKE 1382
 DB 1302 VQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYIYGSSSGGSIHSTLQERCKE 1361
 QY 1383 GEKQEDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPTIETDEKSDGSGKLLPCKKSSSE 1442
 DB 1362 SELQKQESGRKSFMLKRGDVIDYSSSGVSTNEASPLDPTIETDEKSDGSGKLLPCKKSSSE 1421
 QY 1443 RPSLFQTDLLKGGGLRYQKLPDEDESGTGRVOITPHCSNMIRTKLAKAKORCASQOE 1502
 DB 1422 RPSLFQTDLLKGGGLRYQKLPDEDESGTGRVOITPHCSNMIRTKLAKAKORCASQOE 1480
 QY 1503 HSAEPTRTFKAEYLSDALDKDSSDGVSRNNESSPNHSLHNEAADDLSOLEKANLIEL 1562
 DB 1481 HSAEPTRTFKAEYLSDALDKDSSDGVSRNNESSPNHSLHNEAADDLSOLEKANLIEL 1540
 QY 1563 EDEGHSKGRMPHSLSGLOPPIIARMSICSEDKKSPECSLIASSPESWPACOKAYNLN 1622
 DB 1541 EDEGHSKGRMPHSLSGLOPPIIARMSICSEDKKSPECSLIASSPESWPACOKAYNLN 1600
 QY 1623 RTPSTVTNNNTAPTNRANQNDFEIGRETQSQVILRPGSPNPTAVQENLKSMAHKRS 1682
 DB 1601 RTPSTVTNNNTAPTNRANQNDFEIGRETQSQVILRPGSPNPTAVQENLKSMAHKRS 1660
 QY 1683 QRSSTYLSKDSALHHAASSTGFGHERESIL 1715
 DB 1661 QRSSTYLSKDSALHHAASSTGFGHERESIL 1693

RESULT 5
 Q80SX9
 ID Q80SX9 PRELIMINARY; PRT; 1554 AA.
 AC Q80SX9;
 DT 01-JUN-2003 (T=EMBLrel. 24, Created)
 DT 01-JUN-2003 (T=EMBLrel. 24, Last sequence update)


```
Db 1261 TEEDEKSDQSGSKLLPGKKSERPSLFTQDLDLKLKSGSLRYOKLPSDEDESGTESDNTPT- 1319
Qy 1481 CSKMTIRKRLKAKQECASPOEHSABEPIRTIKAKKEYLSDALDKKSDSSGVSRSNESSP 1540
Db 1320 LLKDDKOKKAGSKAERVAKSPHSEVPEIRTIKAKKEYLSDALDKKSDSSGVSRSNESSP 1379
Qy 1541 NLSLHNEAADDSOLEKANLIELEDEGHSGKRGMPHSLSGLQDPIIARMSICSEDKKGPSE 1600
Db 1380 NLSLHNEAADDSOLEKANLIELEDEGHSGKRGMPHSLSGLQDPIIARMSICSEDKKGPSE 1439
Qy 1601 CSLIASSPEESWPACOKAYNLRNPSTVTLNNNTAPTNRANQNDELEGRETQVILRP 1660
Db 1440 CSLIASSPEESWPACOKAYNLRNPSTVTLNNNTAPTNRANQNDELEGRETQVILRP 1499
Qy 1661 GPSNPNTAVQENLKSMAHKESQSSYTRLKSDASELHAASESTGFGEERESIL 1715
Db 1500 GPSNPNTAVQENLKSMAHKESQSSYTRLKSDASELHAASESTGFGEERESIL 1554

RESULT 6
ID Q7T163 PRELIMINARY; PRT; 1680 AA.
AC Q7T163;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SI:d2119J18.2 (Novel protein similar to rat kinase D-interacting
DE substate of 220 kDa (KIDINS220) ).
GN Name:SI:d2119J18.2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL954721; CAEI7588.1; -.
DR HSSP; P42773; 1IHB.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00023; Ank, 11.
DR SMART; SM00248; ANK, 11.
DR PROSITE; PS00088; ANK_REPEAT; 10.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; UNKNOWN_1.
KW ANK repeat; Kinase.
SQ SEQUENCE 1680 AA; 185948 MW; 4FE8AB3377DD5E60 CRC64;

Query Match 70.4%; Score 6233.5; DB 2; Length 1680;
Best Local Similarity 71.8%; Pred. No. 7.3e-315;
Matches 1243; Conservative 180; Mismatches 232; Indels 77; Gaps 26;

Qy 1 MSVLISQVINYVEENIPALKALKERKQVDERNECGQTPLMLAABQGNVIEVKELKN 60
Db 9 MTTLAIQNLFYSYVEENLAAVKVHLDKFEVDGRSDNGQTPMLASQGSLEIQVELIRR 68
Qy 61 GANCNLEDLNWTALISASKEGHIHVEELKSGASLEHRDMGWTALMWACYKGRDVT 120
Db 69 GANYNLDDVDCWSALISAAKEGHVYVVKELIENSAYIEHRDMGWTALTWASYKGRVEA 128
Qy 121 ELLLSHGANSVTLGQYSVVPIIWAAGRGHADIHVLHLLONGAKVNCSDKYGTTPLVWAAR 180
Db 129 TVLLENGANPNNTTQQQSVVPIIWAAGRGHAEIVKLLLEHGAKVNCSDKYGTTPLIWAAR 188
Qy 181 KGHLECVKHLIAGADVDQEGANSMTALIVAVKGGYTSQVKELIKRPNVNLTKDGNNTA 240
Db 189 KGHYDCVHLLLEAGDVDQEGANSMTALIVAVKGGYTEVVKELIKRPNVNLTKDGNNTA 248
Qy 241 LMIASKEGHIIEVDLLDAGTYVNIPIRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
Db 1261 TEEDEKSDQSGSKLLPGKKSERPSLFTQDLDLKLKSGSLRYOKLPSDEDESGTESDNTPT- 1319
Qy 1481 CSKMTIRKRLKAKQECASPOEHSABEPIRTIKAKKEYLSDALDKKSDSSGVSRSNESSP 1540
Db 1320 LLKDDKOKKAGSKAERVAKSPHSEVPEIRTIKAKKEYLSDALDKKSDSSGVSRSNESSP 1379
Qy 1541 NLSLHNEAADDSOLEKANLIELEDEGHSGKRGMPHSLSGLQDPIIARMSICSEDKKGPSE 1600
Db 1380 NLSLHNEAADDSOLEKANLIELEDEGHSGKRGMPHSLSGLQDPIIARMSICSEDKKGPSE 1439
Qy 1601 CSLIASSPEESWPACOKAYNLRNPSTVTLNNNTAPTNRANQNDELEGRETQVILRP 1660
Db 1440 CSLIASSPEESWPACOKAYNLRNPSTVTLNNNTAPTNRANQNDELEGRETQVILRP 1499
Qy 1661 GPSNPNTAVQENLKSMAHKESQSSYTRLKSDASELHAASESTGFGEERESIL 1715
Db 1500 GPSNPNTAVQENLKSMAHKESQSSYTRLKSDASELHAASESTGFGEERESIL 1554
```

```
Db 249 LMTAAKEGYTEIVQDLLDAGTYVNIPIRSGDTVLIGAVRGHVEIVRALLHKYADIDIRG 308
Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLLDKGA 360
Db 309 QENKTALYWAVEKGNATWVRDILQCNPDTEITTKDSETPLIKATKMRNIEVVELLLDKGA 368
Qy 361 KVSADVKKGDTPLHVAIRGRSRRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHQSKIL 420
Db 369 KVSADVKKGDTPLHVAIRGRSRRRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHQSKIL 428
Qy 421 TQIFGARHLSPTEDEGMGLGYDLYSSALADILSPTEWQPPICVGLYQWGGKGFLLKKL 480
Db 429 TQIFGARHLSPTEDEGMGLGYDLYSSALADILSPTEWQPPICVGLYQWGGKGFLLKKL 488
Qy 481 EDEKMTFAGQTEPLFQFWSMLIVPLTLLCGGLGVFAFPVDNTLAIATLSLFLALIYIF 540
Db 489 EDEKMTFAGQTEPLFQFWSMLIVPLTLLCGSVALVIGFTVDPKLAATLSLIALLYVF 548
Qy 541 FVIYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRELF 600
Db 549 FVVYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRELF 608
Qy 601 TDYNRSLSSVGETSLAEIATLSDACEREGFLATRLFRVPRTEESQKKKKKTKCLPS 660
Db 609 TDYNRSLSSVGETSLAEIATLSDACEREGFLATRLFRVPRTEESQKKKKKTKCLPS 668
Qy 661 FVIFLFTVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVLAFVLCNRTWQVLSLL 720
Db 669 FVIFLFTVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVLAFVLCNRTWQVLSVL 728
Qy 721 NSQKRLHSAASKLHLKLSGEGFMKVLKCEVELMARMARKTIDSFQNTQRLVVIIDGLDAC 780
Db 729 NSQKRLHSAANKHMKLSGEGFMKVLKCEVELMARMARKTIDSFQNTQRLVVIIDGLDSC 788
Qy 781 EODKVLQMLDTRVLFSGKGFPIAFASDPHIIKAINQNLSVLNRDNINGHDYMRNIVH 840
Db 789 EODKVLQMLDTRVLFSGKGFPIAFASDPHIIKAINQNLSVLNRDNINGHDYMRNIVH 848
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 849 LPVFLNSRGLSSAKKMCAPAPANGE--TGNSEGHLELDRLKLSQNSLDQTKFGSKTILN 906
Qy 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEDFWDFDISPQTMRRLLNIVSVTGRLLRANQIT 960
Db 907 RRDYRRRQMQRSVTRQMSFDLTKLLVTEDFWDFDISPQTMRRLLNIVSVTGRLLRANQIS 966
Qy 961 FNVDRLASWINLTEQWYRPTSWLLIYLEETEGIPDQMTLKTMYERIKNYPTTKDVEPLL 1020
Db 967 FNVDRLASWINLTEQWYRPTSWLLIYLEETEGIPDQMTLKTMYERIKNYPTTKDVEPLL 1026
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLREIADVRAAREQINIGLAY 1080
Db 1027 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLREIADVRAAREQINIGLAY 1086
Qy 1081 PPLPLHEGPPRPSPGYSQSPASVCS-SASFNGPPPGVSPQPHSSYSGLSGPPQHPFYNR 1139
Db 1087 PTLPLQSG--RPISWYSQSSACSPTASFNGPPYPPGVSQPHSAYFSGMAGPQHPFYNR 1144
Qy 1140 --RAVPATGSSLLISSMTVDVCEKLRQIEGLDONMMPQCTTIKKANINGRVLSCQDID 1197
Db 1145 GSASVSGTFSILLSSMTDVI CERVKLIDGIDQNLSIQYTATIKKANINGRVLSCQDID 1204
Qy 1198 ELKKEMANFGDWHLPFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGSARSSHTEL 1257
Db 1205 ELKKEMANFGDWHLPFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGSARSSHTEL 1262
Qy 1258 PLTESSQTP-YTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRTPSPSSLSNSQSSSIEI 1316
Db 1263 GVAGNTDTSMPYNFLNFSFEELSNVGLLEE-PPRHVNATMGTTHRTPSWSSLSNSQSSSNEI 1321
Qy 1317 SKLTKDQVARYDAYREYIAQMSQLEGGTGSSISGRSSPHSTYIYICQSSGSGSIHSTLE 1376
Db 1322 CKLTKDQVARYDAYREYIAQMSQLEGGTGSSISGRSSPHSTYIYICQSSGSGSIHSTLE 1376
```

QY 1377 QERKGEKQEDGRKSFMLKRG-----DVIDYSSGVSTNEASPLDPTITEDEKSDQS 1430
Db 1368 DKK-KDG--NDQGRKS-VSKRGSTSGSDNTDYASA-----DAATLDPTITEEDKVDHG 1418
QY 1431 GSKLLPGKSS-BRPSLFQ-TDLKLG-GGLRYQKLPSPDESGTGRVQITTPHCSRWIRT 1487
Db 1419 SSKSLGRKTSKGVSLFQADLKLKAGGSRVQKLTSDDEES-----EESDNAPLLKDG 1473
QY 1488 KRLKAKQREKASPOEHSABEPIRPIKAVEYLSALLDKKSSDGSVRSNESSPNHSHNE 1547
Db 1474 KPEAKASDGD-----RSITKGVKLYS-----DKDSSDGSVRSNESSPNHSLQDE 1520
QY 1548 AADSQLEKANLIELEDEGHSRGKRWPHSLSGLODPTIARMSTCSDEKSPSCSLIASS 1607
Db 1521 EADLSOSERANLIEDEESARKRGLPNSLUGLODPTIARMSTCSDE-----QCSLASS 1575
QY 1608 PEESWAPACQAYNLNRTPTVTNNNT--APTNRANQFDEIRET--SOVILRPGPS 1663
Db 1576 PEESWPS-SKSYNLNRTPTVTNNNTNAQQGNHIRQPSDSSNTTSTTTCSDVIINPGTS 1634
QY 1664 PNTAVQNEHLKMAHKSORSRYTLSDKASLHAAASSESTGFBERSIL 1715
Db 1635 TTSATTQENVRVVLKRLGNPG-----DPPBILKVSSETVTFGBERSIL 1690

RESULT 7
Q6P7Y1
ID Q6P7Y1 PRELIMINARY; PRF; 1672 AA.
AC Q6P7Y1;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:63531;
GN ORFNames=zgc:63531;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Srausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061450; AH61450.1;
DR HSSP; Q60773; IAP7.
DR ZFIN; ZDB-GENE-030131-7824; zgc:63531.
DR InterPro; IPR002110; ANK.

DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00023; Ank_10.
DR PRINTS; PR01415; ANKTRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS00088; ANK REPEAT; 10.
DR PROSITE; PS00297; ANK REP REGION; 1.
DR PROSITE; PS00572; GLYCOSYL HYDROL_F1_1; UNKNOWN_1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 1672 AA; 185035 MW; 75BCF10280B410B CRC64;

Query Match 70.3%; Score 6219.5; DB 2; length 1672;
Best Local Similarity 71.7%; Pred. No. 3.8e-314;
Matches 1241; Conservative 180; Mismatches 234; Indels 77; Gaps 26;

QY 1 MSVLISQSVINVEEENIPALKKALKKCKDKVDVERNECGQTPLMLAAEQNGVEIVKLLKN 60
Db 1 MTTLAQNLFVSVEENLAADVHLDPKFEVDGRSDNGQTPLMLASEQSLVQSLIRR 60
QY 61 GANCNLELDNNTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMAWCVKGRDVTV 120
Db 61 GANVNLDVDDCSALISAAKEGHEVEVVKELSENSAYIEHRDMGWTALTWASVYGRVEVA 120
QY 121 ELLSHGANPSVTGLQYSVYPIIWAAGRGHADI VHLLLQNGAKVNCSDKYGTTPLYWAAR 180
Db 121 TVLLENGANPNTTGGQYSVYPIIWAAGRGHABIVKLLLEHGAKEVNCSDKYGTTPLYWAAR 180
QY 181 KGHLECVKHLAMGADVDDGEGANSMTALIVAKGGYTSQSVKELLKKNPNVNLTKDGNTA 240
Db 181 KGHYDCWMLLENGADVDDGEGANSMTALIVAKGGYTSQSVKELLKKNPNVNLTKDGNTA 240
QY 241 LMAISKEGHEIIVQDLIDAGTVNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDRG 300
Db 241 LMAIAKEGTEIVQDLIDAGTVNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDRG 300
QY 301 QNKNTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATKRNTEVVELLDKGA 360
Db 301 QENKNTALYWAKEGNATWVRDILQCNPDTEITTKDSETEPLIKATKRSIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLYRPKNKAGETPNVDCSHQKSL 420
Db 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLYRPKNKAGETPNVDCSHQKSL 420
QY 421 TQIFGARHLSPTEITDGMGLYDLSALADILSEPTMQPPICVGLYVAQSGSKSFLKKKL 480
Db 421 TQIFGARHLSPTEITDGMGLYDLSALADILSEPTMQPPICVGLNTQSGSKSFLKKKL 480
QY 481 EDEMTFAQQQTEPLFPQFSLVIFVTLCCGGLGVFAPVDVTNLAIAISLSFLAIYIF 540
Db 481 EDEMTFAQQQTEPLFPQFSLVIFVTLCCGGLGVFAPVDVTNLAIAISLSFLAIYIF 540
QY 541 FVIVYFGGRREGESNNWAWALSTRLARHIGYLELLEFKLMFVNPPPELPEOTTKALPVRELF 600
Db 541 FVIVYFGGRREGESNNWAWALSTRLARHIGYLELLEFKLMFVNPPPELPEOTTKALPVRELF 600
QY 601 TDYNRLSSVGGTSLAEMIATISDCEREFGLATRLFRVPRTEESQGGKKWKTCCLPS 660
Db 601 TDYNRLSSVGGTSLAEMIATISDCEREFGLATRLFRVPRTEESQGGKKWKTCCLPS 660
QY 661 FVIFLPIVGCITAGITLLAIFRVPDKHLTVNAILISIASVGLAFVLCNRTWQVLDL 720
Db 661 FVIFLPIVGCITAGITLLAIFRVPDKHLTVNAILISIASVGLAFVLCNRTWQVLDL 720
QY 721 NSQKRLHSAASKLHKLKSEGEKVKLCEVELMARMAKTIDSTFQNTOTLVIIIDGLDAC 780
Db 721 NSQKRLHSAASKLHKLKSEGEKVKLCEVELMARMAKTIDSTFQNTOTLVIIIDGLDAC 780
QY 781 EQDKVLQMLDVTVRVLFSGKPFPIAFASDPHIIKAINQNLNSVLRDSNINGHDYMNIVH 840
Db 781 EQDKVLQMLDVTVRVLFSGKPFPIAFASDPHIIKAINQNLNSVLRDSNINGHDYMNIVH 840
QY 841 LPVFLNSRGLSNARPLVTSATNGDITCSDTTGTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARPLVTSATNGDITCSDTTGTGTQEDTDRRVSONSLGEMTKLGSKTALN 900

QY 901 RRDYTRRQMQRTITRQMSFDLTKLLVTFDWFSDISQPMRRLNINIVSVTGRLLRANQIT 960
DB 899 RRDYTRRQMQRSVTRQMSFDLTKLLVTFDWFSDISQPMRRLNINIVSVTGRLLRANQIS 958
QY 961 FNDWELASWINLQTEOWPYRTSWLLIYLBESTEGLPQMTLKWYERISQNIPTTKDVEPLL 1020
DB 959 FNDWELASWINLQTEOWPYRTSWLLIYLBESTEGIPDQTNLKIYERISQNIPTTKDVEPLL 1018
QY 1021 EIDGDIRNFVFLSRTPLVARDVKTEPLCTVNLDPKRLIADVRAAREQINIGGLAY 1080
DB 1019 EIDGVRSEFVFLSRTPLVARDIRTEPLCTVNLDPKRLIADVRAAREQVNMAGVY 1078
QY 1081 PPLPLHEGPRPPSPGSPASVCS--SASFNGFPFGGVSPHSSYYSGLSGPQHPFYNR 1139
DB 1079 PTLPLEG--RPISWYQSSACSPTASFNGFPYPPGVSQPHSAFSGMAGPQHPFYNR 1136
QY 1140 --AAVPAWGSLLSMTVDVCEKROIQGLDQNMPOYCTTIKANINGRVLSQCNID 1197
DB 1137 GSASVSGTSPILLSMTSDVICERVKLIDGIDQNLISQYTATIKANINGRVLSQCNID 1196
QY 1198 ELKKEMANFQDWHLFRSWMLEMSVESQVVPEDPRFLNENSSAPVPHGESARSSHTEL 1257
DB 1197 ELKKEMANFQDWHLFRSWMLEMSVESQVVPEDPRFLNENSSAPVPHGESARSSHTEL 1254
QY 1258 PLTELSSQTP--YTLNFSPEELNTGLDEGAPRHSNLSWQSOTRTPSPSSLSNQSISIEI 1316
DB 1255 VGAGWTDTSPWYFNLSPEELSNVGLER--PPRHVNAWGMATHRTPSNSSLSNQSISNEI 1313
QY 1317 SKLTKDQVAYRDAYREVIAMQSLGEGTGSTISGRSSPHSTHYIYQSSGGGSIHSTLE 1376
DB 1314 CKLTKDQVAYRNAYEDYIASMSQLE-----LGMKPVPP--FVSQ-----LMHSSSE 1359
QY 1377 QERKEGELKQEDGRKSFMLKRG-----DVIDYSSSGVSTNEASPLDPITTEDEKSDQS 1430
DB 1360 DKK--KDG--NPDQGRKS--VSRGSGTSGSDNTDYASA-----DAATLDPITTEDEKVDHG 1410
QY 1431 GSKLLPGKKSS--ERPSLFQ--TDLKL--GGGLRYOKLPSEDESGTGRVQITPHCSKMIRT 1487
DB 1411 SSKSLGRKTSKGVSLFQAGADLKLKAGGSGRYQKLTSDDEES-----EESDNAPLKDQ 1465
QY 1488 KRLKAKORECASPOEHSABPRTTFIKAKYILSDALLDKKSDSGSVRSNESPNSHLSNE 1547
DB 1466 KKPEAKASDGD-----RSLTKGKDYLS-----DKKSDSGSVRSNESPNSHLSQDE 1512
QY 1548 AADPSOLEKANLIELEDECHSGKGMPSHLSGLQDPTIARMSICSEDKKSPSECSLIASS 1607
DB 1513 EADLSQSERANLIELEDENSARKGLPNSLSGLQDPTIARMSICSED-----QCSLIASS 1567
QY 1608 PEESWPACQAYNLNRTPTSTVTLNNT--APTNRANQFDEIGIRET--SQVILRPGPS 1663
DB 1568 PEESWPS--SKSYNLNRTPTSTVTLNNTVAQGNHIRQPSDSNNTSTTTGSDVINPGTS 1626
QY 1664 PNPTAVQENILKMAHKSQRSSTYRLSKDASELHAASSESTGFGEERESIL 1715
DB 1627 TTSATTQENVRVHLKGLNPG-----DPEILKVSSETVTGFEERESIL 1672

RESULT 8

Q9UF42 PRELIMINARY; PRT; 1031 AA.
ID Q9UF42;
AC Q9UF42;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434F0621.
GN Name=DKFZp434F0621;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE-Testis;
RA Blum H., Baurersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (DSC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133620; CAB63746.1; --
DR PIR; T43458; T43458.
DR HSSP; P09959; 1SW6.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS00088; ANK_REPEAT; 10.
DR PROSITE; PS02097; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 1031 AA; 115310 MW; D98866461C13A2F5 CRC64;

Query Match 57.3%; Score 5073.5; DB 2; Length 1031;
Best Local Similarity 95.2%; Pred. No. 6.5e-255;
Matches 974; Conservative 23; Mismatches 13; Indels 13; Gaps 1;

QY 1 MSVLISOSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN 60
DB 1 MSVLISOSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN 60
QY 61 GANCNLEDLNDWTALISASKEGHIHVEELKSGASLEHEDMGWTALMWACYKGRDGV 120
DB 61 GANCNLEDLNDWTALISASKEGHVHVEELKCGVNLHEDMGWTALMWACYKGRDGV 120
QY 121 ELLLSHGANSVTCLOVSVYPIIWAAGRGHADIHLLLONGAKVNCSDKYTTPLWAAAR 180
DB 121 ELLLSHGANSVTCLOVSVYPIIWAAGRGHADIHLLLONGAKVNCSDKYTTPLWAAAR 180
QY 181 KGHLECVKHLAMGADVDDQSGANSMTALIIVAKGGYTSQVKEILKRPNNVNLTDKGN 240
DB 181 KGHLECVKHLAMGADVDDQSGANSMTALIIVAKGGYTSQVKEILKRPNNVNLTDKGN 240
QY 241 LMTASKEGHISIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
DB 241 LMTASKEGHTEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
QY 301 QDNKTALYWAVEKGNATWDRDILQCNPDTEICTKQGETPLIKATKMRNIEVVELLDKGA 360
DB 301 QDNKTALYWAVEKGNATWDRDILQCNPDTEICTKQGETPLIKATKMRNIEVVELLDKGA 360
QY 361 KVSADVKKGTPLHVAIRGRSRLAEILLRNPKDGRLLYRPNKAGETPYNIDCSHQSIL 420
DB 361 KVSADVKKGTPLHVAIRGRSRLAEILLRNPKDGRLLYRPNKAGETPYNIDCSHQSIL 420
QY 421 TQIFGARHLSPETETDGMGLYSSALADILSEPTMQPPICVGLYQWGSKGKSFLLKKL 480
DB 421 TQIFGARHLSPETETDGMGLYSSALADILSEPTMQPPICVGLYQWGSKGKSFLLKKL 480
QY 481 EDEMKTPAGQOTBPLFQFSWLIIVFLTLCLCGGLVFPFVDTNLAIAISLSFALIYIF 540
DB 481 EDEMKTPAGQOIBPLFQFSWLIIVFLTLCLCGGLLFAFTVHPNLGIAVLSFALIYIF 540
QY 541 FIVYFGREGESWMAWALSTELARHICVLELLFKLMFNVPPELPEOTTKALPVPLF 600
DB 541 FIVYFGREGESWMAWALSTELARHICVLELLFKLMFNVPPELPEOTTKALPVPLF 600
QY 601 TDYNLSSVSGETSLEAMIALSDACEREFGLATRLFRVFRTEESQKKWKKTCCPLS 660
DB 601 TDYNLSSVSGETSLEAMIALSDACEREFGLATRLFRVFRTEESQKKWKKTCCPLS 660
QY 661 FVIFLFTVGCIIAGITLLAIFRVDPKHLTVNALISIASVVGAFVLCNCRTWQVLSLL 720
DB 661 FVIFLFTIGCIIISGITLLAIFRVDPKHLTVNALISIASVVGAFVLCNCRTWQVLSLL 720
QY 721 NSQKRLHSAASKLHLKSEGFPMKLVCEVELMARMAKTIDSTQNTQRLWIIDGLDAC 780
DB 721 NSQKRLHNAASKLHLKSEGFPMKLVCEVELMARMAKTIDSTQNTQRLWIIDGLDAC 780
QY 781 EQKVLQMLDTRVFLFSKGPFIAPASDPHIIKAINQNLNSVLNRDSNINGHDMRNVH 840

```
Db 781 EQDKVLQMLDVTVRVLSKGPPIAFASDPHIIKAINQNLNSVLRDSNINNGHDMYRNIVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSPTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDSTTGTQEDADRVSQNSLGEKTKLGSKTALN 900
QY 901 RRDYRRRQWRITITRQMSFDLTKLVTEWDFSDISPTQMRRLNLTIVSVTGRLLRANQIT 960
Db 901 RRDYRRRQWRITITRQMSFDLTKLVTEWDFSDISPTQMRRLNLTIVSVTGRLLRANQIS 960
QY 961 FNDRLASWINLAEQMPYRTSWLILYLETEGLPDQWTLKTYER-----IS 1007
Db 961 FNDRLASWINLAEQMPYRTSWLILYLETEGLPDQWTLKTYERCCGADSCDRDRIGIS 1020
QY 1008 KNI 1010
Db 1021 KSV 1023

RESULT 9
Q72322 PRELIMINARY; PRT; 691 AA.
AC Q72322;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686H14204 (Fragment).
GN Name=DKFZp686H14204;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human fetal brain;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BX538189; CAD98059.1; -
DR InterPro: IPR010993; SAM_homology.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 691 AA; 75760 MW; F5783F24DDEA6027 CRC64;

Query Match 35.08; Score 3097; DB 2; Length 691;
Best Local Similarity 86.8%; Pred. No. 1.8e-152;
Matches 601; Conservative 32; Mismatches 57; Indels 2; Gaps 2;

QY 1025 DIRFEVFLSSRTPVLVARVKTFLPCTVNLDPKLRREIIADVRAAREQINIGGLAYPPLP 1084
Db 1 DIRFEVFLSSRTPVLVARVKTFLPCTVNLDPKLRREIIADVRAAREQISIGGLAYPPLP 60
QY 1085 LHGGPPRPPGSGYOPASVCSSAFNGFPFGVSPQPHSSYSGLSGPGQHPFFNRAAVPA 1144
Db 61 LHGGPPRAPSGYOPSPSCSTSFNGFPAGVSPQPHSSYSGMTGPPQHPFFNRRSGPA 120
QY 1145 TGSLLSSMTVDVVCCKRQIEGLDQNMMPQVCTTIKKNANINGRVLSCQNIDELKEMA 1204
Db 121 PGPVLLNSLVDAVCEKQKQIEGLDQSMPLQVCTTIKKNANINGRVLQACNIDELKEMN 180
QY 1205 MNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTELPTELSS 1264
Db 181 MNFGDWHLFRSTVLEMEANESHVVPEPRFLSESSGPA PHGPARRASHNELPHTLSS 240
QY 1265 QTPYTLNFSFEELNTLGLDGAAPHSLNSWQSOTRRTPLSLNSQDSSIEISKLDKQV 1324
Db 241 QTPYTLNFSFEELNTLGLDGAAPHSLNSWQSOTRRTPLSLNSQDSSIEISKLDKQV 300
QY 1325 AEYRDAYREYIAQMSOLEGGTGSSTISGRSSPHSTYIGOSSGGSTHSTLEOERKEGE 1384
Db 301 AEYRDAYREYIAQMSOLEGGPGSTTISGRSSPHSTYIGOSSGGSTHSTLEOERKQDSE 360
QY 1385 LKQEDGRKSLMKRGDVIDYSSSGVSTNEASPLDPITEDEKQDQSGSKLLPGKKSERP 1444
```

```
Db 361 PXPDDGRKSLMKRGDVIDYSSSGVSTNEASPLDPITEDEKQDQSGSKLLPGKKSERS 420
QY 1445 SLFQTDLKLKGGGLRYQKLPSPDESGTGRVQITHCSSKMTKRLKAKORECASQEHHS 1504
Db 421 SLFQTDLKLKGGGLRYQKLPSPDESGTGRVQITHCSSKMTKRLKAKORECASQEHHS 479
QY 1505 ASPIRTFIKAKVLSDALDKKSDSGVRSNESPNSHNEAADDSOLEKANLIELED 1564
Db 480 ASPIRTFIKAKVLSDALDKKSDSGVRSNESPNSHNEAADDSOLEKANLIELED 539
QY 1565 EGHSGRGRGMPHSLSGIQLQDPPIIARMSICSEDKSPSECSLIASSPEESWPACQKAYNLRT 1624
Db 540 DSHSGRGRGMPHSLSGIQLQDPPIIARMSICSEDKSPSECSLIASSPEESWPACQKAYNLRT 599
QY 1625 PSTVTLNNTAPTNRANQNFDIEGIRETSQVILRPGSPNPTAVQENLKSMAHKRSOR 1684
Db 600 PSTVTLNNTAPTNRANQNFDIEGIRETSQVILRPGSPNPTAVQENLKSMAHKRSOR 659
QY 1685 SSYTRLSKDASELH-AASSESTGFGERESIL 1715
Db 660 SSYTRLSKDPPPELHAAASSESTGFGERESIL 691

RESULT 10
Q6MZU2 PRELIMINARY; PRT; 664 AA.
AC Q6MZU2;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686A19189 (Fragment).
GN Name=DKFZp686A19189;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human esophagus tumor;
RA The German Human cDNA Consortium;
RA Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BX540878; CAE45935.1; -
DR InterPro: IPR010993; SAM_homology.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 664 AA; 72879 MW; 678BA3EF27DA1BDF CRC64;

Query Match 30.9%; Score 2732; DB 2; Length 664;
Best Local Similarity 80.6%; Pred. No. 1.6e-133;
Matches 536; Conservative 31; Mismatches 58; Indels 40; Gaps 3;

QY 1090 PRPPGYSOPASVCSSAFNGFPFGVSPQPHSSYSGLSGPGQHPFFYNR----- 1139
Db 1 PRAPSGYOPSPSCSTSFNGFPAGVSPQPHSSYSGMTGPPQHPFFYNRFFAPLYTP 60
QY 1140 -----AAVPATGSSLLSSMTVDVVCCKRQIEGLDQ 1171
Db 61 RYYPGSGQHLISRPVSKTSLPRQNNGLSGGPA GPVVLNLSLVDAVCEKQKQIEGLDQ 120
QY 1172 NMMPOCTTIKKNANINGRVLSCQNIDELKEMANFQDWHLFRSMVLEMRSVESQVVPED 1231
Db 121 SMLPQCTTIKKNANINGRVLQACNIDELKEMANFQDWHLFRSTVLEMEANESHVVPE 180
QY 1232 PRFLNENSSAPVPHGESARRSHTELPTELSSQTPYTLNFSFEELNTLGLDGAAPHSN 1291
Db 181 PRFLSESSGPA PHGPARRASHNELPHTLSSQTPYTLNFSFEELNTLGLDGAAPHSN 240
QY 1292 LSWQSOTRTPSLSLNSQDSSIEISKLDKQVQAEYRDAYREYIAQMSOLEGGTGSSTIS 1351
Db 241 LSWQSOTRTPSLSLNSQDSSIEISKLDKQVQAEYRDAYREYIAQMSOLEGGTGSSTIS 300
```

```
QY 1352 CRSSPHSTYYICQSSSGGSIHSTLEQERKGEKQEDGRKSFMLKRGDVIDYSSSGVST 1411
DB 301 CRSSPHSTYYMGQSSSGGSIHNSLEQEKGDSEPKPDGRKSFMLKRGDVIDYSSSGVST 360
QY 1412 NEASPLDITTEDEKSDSGSKLLPGKSSSRPSLFQTDLKGGLRYQKLPDEDSG 1471
DB 361 NDASPLDITTEDEKSDSGSKLLPGKSSSRPSLFQTDLKGGLRYQKLPDEDSG 420
QY 1472 TGRVQITPHCSKMTITKELKAKQREKSCASPFQBSHAPITFTIKAKYELSDALLDKKSDS 1531
DB 421 TEESDNTPLKDKDKRAEKGKVERVPKSPHSAEPITFTIKAKYELSDALLDKKSDS 479
QY 1532 GVRNESPNNHNEAADDQLEKANLIELEDEGHSGKRGKMPHSLSGLOPPIIARMSIC 1591
DB 480 GVRSESPNNHNEAADDHLEKANLIELEDDSHSGKRGKPHSLSGLOPPIIARMSIC 539
QY 1592 SEDKKSPECSLIASSPESPACOKAYNLARTSTVTNLNNTAPTNRANQNFDEIEGIR 1651
DB 540 SEDKKSPECSLIASSPESPACOKAYNLARTSTVTNLNNTAPTNRANQNFDEIEGIR 599
QY 1652 ETSQVILRPSPNPTAVONENLMSMAHRSORSSYTRLSKDASLH-AAASESTGFGE 1710
DB 600 ETSQVILRPSPNPTAVONENLMSMAHRSORSSYTRLSKDASLH-AAASESTGFGE 659
QY 1711 RESIL 1715
DB 660 RESIL 664
QY 1711 RESIL 1715
DB 660 RESIL 664
RESULT 11
Q7KVP5 PRELIMINARY; PRT; 1604 AA.
AC Q7KVP5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE CG30387-PC.
GN ORFNames=CG30387;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
```

```
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel N.S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.,
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
[5]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003453; AAM70867.2; -
DR HSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 10.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 12.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 1604 AA; 176580 MW; 3259E2F4F183229A CRC64;
Query Match 27.4%; Score 2421.5; DB 2; Length 1604;
Best Local Similarity 35.8%; Pred. No. 8.7e-117;
Matches 584; Conservative 297; Mismatches 585; Indels 165; Gaps 42;
QY 1 MSVLISQSVINYVEENIPALKALLE-KCKDVEDNECGOTPLMAEEQGVNVEIKLLK 59
DB 1 MGSIGHRALLOQYIDNNDISGLRAILDSRLHITDDRDNATTVLMVAVRGLTAFVRBFLA 60
```


QY 1044 DVKTELPCTVNDPKLEIITADYRAAEQINIGLAVPPLPLHEGP-----PRPPSGYSQP 1099
DB 1086 DLRIFFPFTINLPYRKLUKEQOQTIE--DEGSLVQIARPSVNTWRQPPATTVVPS 1143
QY 1100 ASVCSSAFNGPPGGVVSQPHSSVYSGLSGQHPFYNAAPATCGSSLL---LSSMTV 1156
DB 1144 QAPPPYQMFONEPAN---ELSRNLSTSTETVPLINSFS--DSFGDDILOTKLTDLT 1198
QY 1157 DVVCEKLRQIEGLDQNNMPOYCTTIKANINGRVLSCNTIDELKEMANFGDWHFRSM 1216
DB 1199 EGVISLDRIDEDM-KPALPKLAPVLRENAINGRVLKCHDMPDLKSVGLSGFHWELFRLL 1257
QY 1217 VLEWRSVE-----SOVVPDPRLNENSSAPV-----PHGESARS---SHTPLP 1258
DB 1258 ITTLRECELRPRKQOQOQOQGALEAPSNVPMIKDVTDALMQPFRSLRSKNSVSHMEKQ 1317
QY 1259 LT---ELSSQTPYTLNFSPELNTLGLDEGAPRHSNLSWOSQTRRTPSLSSL-----N 1308
DB 1318 VTLEEQMICGTLQTLNEEAYE-DVASSERSPTGEMLAQAOLQAPIRESSEFGSPSD 1376
QY 1309 SQSSSEISIKLTKVQ---AEY-RDAYREYIAQMSQI-----EGGTGSGSTI---SGRSS 1355
DB 1377 QKQYGVKISNNNNNNNYLHAEYRNSVSSHLSQSLTLVGA PVGHGGGGSHLHGNGNDL 1436
QY 1356 PHSTY-----YIGSSSG-----SIHSTLEQER-----GKEGE 1384
DB 1437 SDSTLDMHVDVSFVGGGGGYHRASQISISELHESKLDTPSAGAAWTPPLGASGS 1496
QY 1385 LKQEDGRKSPMKRGDV-----IDYSSSGVSTNEASPLDP-----ITREDEKSPQSGS 1432
DB 1497 VPAPSGRESLTKQGSVKADKRVSIQOMATSNNNNNSTKLTNVEVSRQEVQAGK 1556
QY 1433 KLPLPGKSSRPS---LFQTDLKLGGGLRYQKLPDSDESGTGRVQITPHCSKMIRTKR 1489
DB 1557 RLATKPPGPRPASLIITRDNSSQFOLLRSSVYDVEDVAQEHRT-----TIRTTL 1608
QY 1490 LKAKQRECAP 1500
DB 1609 LEQOEESAP 1619
RESULT 13
Q9W210 PRELIMINARY; PRT; 1678 AA.
AC Q9W210;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE CG30387-PB.
GN ORPNames-CG30387;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,

Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
Rostler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodsk A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195 (2000).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RN Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RN Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003453; AAF46711.4; --
DR HSSP; P42773; 1IHB.
DR IntAct; Q9W210; --
DR FlyBase; FBgn0050387; CG30387.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 10.
DR PRINTS; PR01415; ANKYRIN.

DR SMART; SM00248; ANK; 12.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK_REPEAT.
SQ SEQUENCE 1678 AA; 184516 MW; D75A23F12E545DFC CRC64;

Query Match 27.4%; Score 2421.5; DB 2; Length 1678;
Best Local Similarity 35.8%; Pred. NO. 9.3e-117;
Matches 584; Conservative 297; Mismatches 585; Indels 165; Gaps 42;

QY 1 MSVLISQSVINVEENIPALKALLE-KCKDVERNECGQTPLMLAEOGQNVVEIVKELLK 59
DB 75 MSLGRLALQYIDNNDISGRALDLSRHLTDIDDENATTVLMVVGAGLTAFAVREPLA 134
QY 60 NGANCNLEDLNTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMACYKGRDVI 119
DB 135 RGADVQAEEDLNTALLCASRNGHLDVVQLLDHGAVERHDMGWTSLMAAARGHTEL 194
QY 120 VELLISHGANFVTLQSVPIIWAAGRHADIVHLLQLONGAKVNCSDKYTGTTPLVWAA 179
DB 195 VRLLLDKGADGNAHG-NYHLGALLWAAGRYKDIVELLVQRGAKVNVGDKYTGTTALVWAC 253
QY 180 RKHLECVKHLKACADVQDSGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDQNT 239
DB 254 RRGVNEIVDTLLKAGANVDTAGMYSWTPLLVAAAGGHTDCVSSILEKKNVNALDKDGMT 313
QY 240 ALMIASKEGHIIEIVQDLILDAGTYVNIIDRSGDTVLIGVRGCHVEIVRALLQKVDIDIR 299
DB 314 ALCTASREGFODIAASLIAAGAYINIQDAGDTPLIHAVKAGHTVVEALLKKGADVDIQ 373
QY 300 GODNKTALYVAEKGNAWMDRILOCPNDBTEICTKOGETPLIKATKRNIEVVEVLLLDKG 359
DB 374 GKDRKTAITYAVEKGHTPIVLLATNPDESATKOGDTPLLRVNRNLEIVHLLDRK 433
QY 360 AKVSAVCKGDTPLHVAIRGRSRRLAELLINPKDGRLLYRPNKAGETPPYINIDCHOKSI 419
DB 434 AKVTASDKRGDTCHIAWRASKTIVETALLRNKHSQLLYRANKAGETPPYINIDSLHOKTI 493
QY 420 LTQIFGARHLSPTETDGMGLDYLYSSALADILSEPTMOPPICVGLVAOAGSGKSFLLKK 479
DB 494 LGQVFGARRLNTNDESGMLGYLYSSALADVLSEPTITPTITVGLVAKMGSGKSFLLNK 553
QY 480 LEDEMKTFAGQOTPLFQFWSMLIVFLTLGLCGGLVPFAPV-DTNLAIAISLSFLALIY 538
DB 554 LRDEMNFAQWAPPIRTSGLLFVCLHVALLLIGTIVGLTSVAVGVSAVGFLLAY 613
QY 539 IFFVIVFGREGESNNWALSTRLARHIGYLELLFKLMFVNPDELPEQTTVALPVRF 598
DB 614 LLLAAVRYCNYQ--MDMQWAVSVQGLEKRMTRLRLLIQVAFCHPPG-POSQSQAKPVRF 670
QY 599 LFTDYNRLSSVGGTSLAEMIATLSACERFGLATRLPRVPRTE--ESQKKKKKTC 656
DB 671 HFAEANSAPT-GOGVAHMLAALLDAIESHYGWLATRLYRFPKCLKVDVGVWRWRMC 729
QY 657 CLPSFVIFLPIVGCIIAGITLAIIF-----RVDPKHLTVNAILISIASVVGGLAFVLCRT 711
DB 730 CIPVILFELALVTVTGTSITVAYFTFADEKEKEHILV--ALVIAAVGTLCTHLHV 787
QY 712 WQVLDLSLMSQRLHSAKHLKSEGMKVLKCEVELMARMATIDSTFQNTQRLV 771
DB 788 LAKVFLVFLTSRHLVLRKAV-----RSSESAPLTLMGAEVAVMTDMVKCLDAFTNQQRSLV 843
QY 772 VIIDGLDACQDKVLQMLDVTVRVLF--KGPPIAFASDPHILIKAINQNLNSVLRDSNI 829
DB 844 GVIDALSDCTERLTLTINAVQTLLSPNRPVLLISVDPHIVIAKAAANSRLFTGGI 903
QY 830 NGHDMYRNIVHLPVFLNSRGSNARKFLVTS-----ATNGDITGSD--TTGTQDQDTRRV 882
DB 904 GGHDPLNLVHLPYQLNSGLRKVQRAQMTALLPKRSGGQDYQDDQPTLG-HSVSARLL 962
QY 883 SONS--LGEMTKL-----GSKTALNR-----RDTVRRRQMTITRQMSFDITKL 925
DB 963 SNASEIISQBLKRGAPARGGGKKLRLESVASSTGSLNHLRGLGNPQTV-----LDLSRI 1017

QY 926 LVTEDFWSDISPTQMRLLNIVSVTGRLLRANQITFNWDRLASMINLTEQWPYRTSWIL 985
DB 1018 VLTDDYFSDVNPVRSMRLMNVYITVLLKAFQIEFYSWYRLSSNINLTEQPLASMIVL 1077
QY 986 YLEB--TEGLPDQMTLKMTERISKNIPTTKDVPLEIDGDIRNFVFLSSRTPVLVAR 1043
DB 1078 HHQFMDSNADESVSQSVVEKLPKLAFLREAAAPLLELDRDKLDAFLQLHKSDDLVA 1137
QY 1044 DVKTFLPCTVNLDPKLEIITADVRAAEQINIGLAVPPLHGEPP-----PRPSPGYSQP 1099
DB 1138 DLRIPLFTNLDPYLRKVLKEDQQTIE--DEGSLVIOARPSVNTMRQFPAPTYVPS 1195
QY 1100 ASVCSSASFNGPPFGVVPQPHSYYSGLSGPQHPPFNRAAVPATGSSLL---LSSMTV 1156
DB 1196 QAYPPYQMFQNEYAN---ELRSNLSTSTPTVPLINSPS-DSFGDDILQTLKLTDTV 1250
QY 1157 DVCEKLRQIEGLDQNMPOYCTTIKKANINGRVLSQCNIDELKKEMANFGDMHFRSM 1216
DB 1251 EGVISLLDRIDEM-KPALPKLAPVLRNAINRVLKHCMDPLKSVLGLSGHWEPLRL 1309
QY 1217 VLEMRSE-----SOVVPEDPRFLNENSAVP-----PHGESARRS---SHTPL 1258
DB 1310 ITTLRECELRPRKQROQOQPGALEAPSNVPMIKDVTDALMQPPRESLRKXNSVSHMEKQ 1369
QY 1259 LT---ELSSOTPYTLNFSFEELNTLGLDEGAPRHSNLSWOSOTRRTTSLSSL-----N 1308
DB 1370 VTLEEQMICGLTQLNEEAYE-DVASSERSPTCEMLAAVAQLQAFIRESEFSGSDD 1428
QY 1309 SQSSSIBSKLTDKVQ---AEY-RDAYREYIAQMSQL-----EGGTGSSTI---SGRSS 1355
DB 1429 QKQYGVKISNNNNNYLHAEYRNSVSSHLSQSLSTLVGAPVGHGGSGSHLGNNDL 1488
QY 1356 PHSTY-----YIGQSSSG-----SIHSTLEQER-----GKEGE 1384
DB 1489 SDSTLDLMHVDSVFGGGGGYHRASQISISELLHESKLDTPSAGAAMTTPLLGASGS 1548
QY 1385 LKQEDGRKSLMKRGDV-----IDYSSGVSTNEASPLDP---ITEDEKSDSGS 1432
DB 1549 VPAPSGRESILKQGSVKADKRVSIQOMATSNNNNNSTKLTNPVVEVSERQPEVQAGK 1608
QY 1433 KLLPCKKSSRPS---LFQTDLKLKGLRLYQKLPSEDESGTGRVQITPHCSKMTKR 1489
DB 1609 RLTTKPPGPRPASLIITRDNSSNQFLLRSSVDYDDVEAQEHRT-----TIRTL 1660
QY 1490 LKAKQRECAP 1500
DB 1661 LEQOEESAP 1671

RESULT 14
QY9H94 ID QY9H94 PRELIMINARY; PRT; 543 AA.
AC QY9H94
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein FLJ12811.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Oeuchi T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Negahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,


```
Db 542 YFTDILLKFLDR--YDLEWYSLNYGLSRKLRLLQLQVAFCHPPG--POSDDQMPVR 598
Qy 598 FLFTDYNRSLSGVGETSLAEMIATLSDACEREFGLATRLPRVERTE--BSQKKKKWKT 655
Db 599 FHFAEASG--AAPGDAVALMLASLFDIAEIAHYGSLATGLYRAFPLKATGGKWKRM 657
Qy 656 CCLPSFVIFLPIVIGIAGITLLAIFRVPDKHLTVNA---ILISIASVVGAP---VLNC 709
Db 658 CCMF--VVLMEFELGML--GILATASLSIVSYSEYGLEGREIAVAIYVLLGILLAGATNL 713
Qy 710 RTWQVLDLSLNSQKRLHSAASKLHLKLSGPFMKVLKCEVELMARMAKTIDSFTQNT 769
Db 714 HAWSKLIALGALFMSOGKHLKRAFNS-----NEAAPLTALGAEVSMLTDMVRCILDAFTGQOSR 769
Qy 770 LVWIIDGLDACQDKVQLMDTVLPLPS--KGPFIATFASDPHIIKAINONLNSVLKDS 827
Db 770 LVGVVDALDSCDTERLTILNAVQTLISGPORPPFVLLAVDPHVIAXAAEANSRRLFT 829
Qy 828 NINGHDYMRNIVLHPVFLNSRGLSNARKPLVTSATNGDITCSDTTGTQEDTD----- 879
Db 830 GIGGHDFLRNLVHLPLVYQLNSGLRKVKVQRAQNTAMSAVRRAMPDLS---RDDQPHLG 886
Qy 880 --RRVSQNS--LGEMTKLGSKTALNRDITYRRQWQ-----RTITRQMSFDLTKL 925
Db 887 SVRSLNASEIMSQEKLAMPSSSRGSKLRVSDSIASSIGNLHKLQGNPPVDLSKI 946
Qy 926 LVTEDFWSDISPOTRRLNLIVSVTGRLLRANOITFNWDLASWINLTQOWPYETSMIL 985
Db 947 ILTDYFSDVNPMSRRRLNVIYITVRLKAFQIDFSWYRLSSWINLTQOWPLRASMI 1005
Qy 986 YLEETEG---LPDQMTLKTMYERISKNIPTTKDVEPLLEIDGDIRNFEVFLSSRTP 1042
Db 1006 -LEHDQGDSPDSDSVLSQAVYDKVRPKLTCLEAAALDLDLDRDKLDAFLQLHKS 1064
Qy 1043 RDVTFPLCTVNLDPKUREIIADVRAAREQINIGGLAYPPLPLHEGPPRPPSGYSQ 1102
Db 1065 SDRIFLFPFTNLDPYLRKVLKEDQALE--DEGGVGTNALMGEPLPKP----- 1112
Qy 1103 CSSASFNGPPPGVVSQPHSSYVYSGLSGPGQHPFVNRAAVATG----- 1146
Db 1113 ---GSVGG---GPIITEHPHEQHSITLGSNS-----GRSSKTTTALOKSSNGLAT 1162
Qy 1147 ----SLLLSMTYDVVCEKLRQIEGLDQNNMPOYCTTIKKANINGRVLSQCNI 1202
Db 1163 DIDLSNQLSSLTVEQLIELEKVNDL-KPAWERTAPILRENAISGRVLMYCNLEEL 1221
Qy 1203 MAMNFGDWHLPRSVLEWRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTPLTE 1261
Db 1222 LRLSFGHWMFKLLVLTALR--EASVTPASRKLSTKIT-----SPAKGTNDSVEN 1272
Qy 1262 LSSQTPYTLNPSFBEALNTLGLDEGAPRHSNLSWQSOTRTPSLSLNSQDSSIEISK 1321
Db 1273 IAQSTP-----PFGSSTS-SFQPIRQKSNLLEKQ-VTLEEQMICG 1311
Qy 1322 KVQAEYRDVAREYIAQMSQLEGGTSGSTISGRSSPHSTY-----IGQSSSGS 1370
Db 1312 ALQTLNEDAFEDVYSS-----SERPSTGEMFSQYLAPIRESSEIGS 1353
```

Search completed: February 5, 2005, 19:33:57
Job time : 198.5 secs

This Page Blank (uspto)